UNITEXT 157

Pablo Pedregal

Functional Analysis, Sobolev Spaces, and Calculus of Variations

Springer

UNITEXT

La Matematica per il 3+2

Volume 157

The **UNITEXT - La Matematica per il 3+2** series is designed for undergraduate and graduate academic courses, and also includes books addressed to PhD students in mathematics, presented at a sufficiently general and advanced level so that the student or scholar interested in a more specific theme would get the necessary background to explore it.

Originally released in Italian, the series now publishes textbooks in English addressed to students in mathematics worldwide.

Some of the most successful books in the series have evolved through several editions, adapting to the evolution of teaching curricula.

Submissions must include at least 3 sample chapters, a table of contents, and a preface outlining the aims and scope of the book, how the book fits in with the current literature, and which courses the book is suitable for.

For any further information, please contact the Editor at Springer: francesca.bonadei@springer.com

THE SERIES IS INDEXED IN SCOPUS

***

UNITEXT is glad to announce a new series of free webinars and interviews handled by the Board members, who rotate in order to interview top experts in their field.

Access this link to subscribe to the events:

https://cassyni.com/s/springer-unitext

Pablo Pedregal

Functional Analysis, Sobolev Spaces, and Calculus of Variations

Springer

Pablo Pedregal
Department of Mathematics
University of Castilla-La Mancha
Ciudad Real, Spain

ISSN 2038-5714 ISSN 2532-3318 (electronic)
UNITEXT
ISSN 2038-5722 ISSN 2038-5757 (electronic)
La Matematica per il 3+2
ISBN 978-3-031-49245-7 ISBN 978-3-031-49246-4 (eBook)
https://doi.org/10.1007/978-3-031-49246-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

To María José, for being always there

Preface

This is a textbook that is born with the objective of providing a first, basic reference for three important areas of Applied Analysis that are difficult to separate because of their many interconnections; yet they have their own personality, and each has grown to become an independent area of very live research. Those three areas occur in the title of the book. There is, at least, a fourth one, Partial Differential Equations (PDEs), which is also quite close to those other three but has such a strong temperament, and attract so much attention, that there are many popular and fundamental references to cover this field in connection with the others. Though there is no reasonable way to avoid talking about PDEs in this text, our intention is to focus on variational methods and techniques, since there are many more references to cover PDEs. In particular, the popular textbook [Br11] has been an inspiration for us.

We should also warn readers from the very beginning that our variational techniques are more oriented towards the direct method rather than to other techniques that could be placed under the umbrella of indirect methods, meaning by this all fundamental results related to optimality. We will of course cover too the Euler-Lagrange equation, but we will not elaborate on this in any special manner. As a matter of fact, some of these ideas are very hard to extend to the higher dimensional situation. On the other hand, the material on Sobolev spaces covered is also oriented and motivated by the needs of variational problems, and hence some basic facts about these important spaces of functions are not proved or treated here, though they can be very easily shown once readers mature the basics.

As such, this book can serve several purposes according to where emphasis is placed. Ideally, and this is how it was conceived, it is to be used as a first course in the Calculus of Variations without assuming that students have been previously exposed to Functional Analysis or Sobolev spaces. It is therefore developed with no prerequisites on those basic areas of Analysis. But it can also serve for a first contact with Functional Analysis and one of its main applications to variational problems. It may also be used for a first approach to weak derivatives and Sobolev spaces, and its main application to variational problems and PDEs. More specifically:

- A first course in Calculus of Variations assuming basic knowledge of Functional Analysis and Sobolev spaces: Chaps. 3, 4, 8, 9.
- A basic course in Functional Analysis: Chaps. 1, 2, 3, 5, 6.
- An introductory course on Sobolev spaces: Chaps. 1, 2, 3, 7.

Pretending that there are no other areas of Applied Analysis involved in such an endeavor would be narrow-minded. To name just one important example, there is also abundant references to the classic subject of Convex Analysis as treated in very classic sources like [EkTe99] or [Ro97]. Even within the field of the Calculus of Variations, we do not aim at including all that is to be learnt. There is, for instance, a lot of material more related to the indirect method that we hardly treat. All of this is very well represented in the encyclopedic work [GiHil96]. The final Appendix is an attempt both to enlarge the perspective of students after covering this material, and to describe, in a non-exhaustive way, those many other areas of Analysis directly or indirectly related to variational problems.

The book is addressed to students and new comers to this broad area of optimization in the continuous, infinite-dimensional case. Depending on needs and allotted time, the first part of the book may already serve as a nice way to taste the three main subjects: Functional Analysis, Sobolev spaces and Calculus of Variations, albeit in a one-dimensional situation. This part includes the basic notions of Banach spaces, Lebesgue spaces and their primary properties; one-dimensional weak derivatives and Sobolev spaces, the dual space, weak compactness, Hilbert spaces and their basic fundamental properties; the Lax-Milgram lemma, the Hahn-Banach theorem in its various forms, and a short introduction to Convex Analysis and the fundamental concept of convexity. The final chapter of this first part deals with one-dimensional variational problems and insist on convexity as the fundamental, unavoidable structural assumption. The second part focuses on some of the paradigmatic results in Functional Analysis related to operators, understood as mappings between infinite-dimensional spaces. Classic results like the Banach-Steinhauss principle, the open mapping and closed graph theorems, and the standard concepts of linear continuous operators in parallelism with the usual concepts of Linear Algebra are examined. The important class of compact operators are dealt with in the second chapter of this part. Finally, the third part focuses on high-dimensional problems: we start introducing multi-dimensional Sobolev spaces and their basic properties to use them in scalar, multi-dimensional variational problems; a presentation of the main facts for this class of problems follows; and a final chapter pretends to cover more specialized material where students are exposed to the subtleties of the high-dimension case. At any rate, each chapter is like a first introduction to the particular field that is treated. Even chapters devoted to variational techniques are far from being complete in any sense, so that readers interested in more information or in deepening their knowledge should look for additional resources.

The scope of the subject of this book is so broad that the relevant literature is truly overwhelming and unaffordable in finite time. Being a source addressed to students

with no previous background on these areas, references are restricted to textbooks, avoiding more specialized bibliography.

The style of the book is definitely pedagogical in that a special effort is made in conveying to students the reasons of why things are done the way they are. Steps ahead searching for new concepts are tried to be justified on the grounds that they are necessary to advance in the understanding of new challenges. Proofs of central results are however a bit technical at times, but that is also a main part of the training that is sought by the text. There is a good collection of exercises, aimed at helping to mature the ideas, techniques and concepts. It is not particularly abundant to avoid that frustration feeling of not going through them all, but sufficient to that purpose. Some exercises are taken from certain steps in proofs. Thinking about a semester, there is no enough time to go through many more exercises. Additional ones can be found in the bibliography provided.

Prerequisites include, in addition to Linear Algebra, Multivariate Calculus, and Differential Equations, a good background on more advanced areas such as Measure Theory or Topology. A partial list of fundamental facts from these areas that are used explicitly in some proofs are:

1. The multivariate integration-by-parts formula.
2. Measure Theory: dominated convergence theorem, Fatou's lemma, approximation of measurable sets by open and compact sets.
3. Lebesgue measure in $\mathbb{R}^N$.
4. Fubini's theorem.
5. Arzelá-Ascoli theorem.
6. Radon-Nykodim theorem.
7. Tychonoff's theorem on compact sets in a product space.
8. Urysohn's lemma and Tietze extension theorem.
9. Egorov's and Luzin's theorems.

A final remark is worth stating. Since our emphasis in this book is on variational methods and techniques, whenever possible we have given proofs of results of a variational flavor, and hence such proofs are sometimes not the standard ones that can be found in other sources.

I would like to thank the members of the Editorial Board of the series UNITEXT for their support; my gratitude goes especially to Francesca Bonadei and Francesca Ferrari for their constant assistance.

Ciudad Real, Spain — Pablo Pedregal
November 2023

Contents

Chapter 1
Motivation and Perspective

1.1 Some Finite-Dimensional Examples

Most likely, our readers will already have some experience with optimization problems of some kind, either from Advanced Calculus courses or even from some exposure to Mathematical Programming. The following are some typical examples.

1. Minimizing the distance to a set. Let

$$\rho(x, y) = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

be the distance function to a given point

$$\mathbf{P} = (x_0, y_0) \in \mathbb{R}^2.$$

Suppose we are given a set $\Omega \subset \mathbb{R}^2$ with $(x_0, y_0) \notin \Omega$, and would like to find the closest point to $\mathbf{P}$ in Ω. With some experimentation for different types of sets Ω, can our readers conclude under what conditions on Ω one can ensure that there is one and only one point in Ω which is closest to $\mathbf{P}$? Does the situation change in three-dimensional space $\mathbb{R}^3$, or in $\mathbb{R}^N$, no matter how big N is?
2. Assume we have two moving points in the plane. The first one is moving in the X-axis from right to left, starting at $x_0 > 0$ with velocity $-u_0$, $u_0 > 0$; while the second is moving vertically from bottom to top, starting at $y_0 < 0$ with speed $v_0 > 0$. When will those objects be the closest to each other? What will the distance between them be at that moment?
3. A thief, after committing a robbery in a jewelry, hides in a park which is the convex hull of the four points

$$(0, 0), \quad (-1, 1), \quad (1, 3), \quad (2, 1).$$

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*, La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4_1

The police, looking after him, gets into the park and organizes the search according to the function

$$\rho(x_1, x_2) = x_1 - 3x_2 + 10$$

indicating density of surveillance. Recommend the thief the best point through which he can escape, or the best point where he can stay hidden.

4. Divide a positive number a into n parts in such a way that the sum of the corresponding squares be minimal. Use this fact to prove the inequality

$$\left(\sum_{i=1}^{n} \frac{x_i}{n}\right)^2 \leq \sum_{i=1}^{n} \frac{x_i^2}{n}$$

for an arbitrary vector

$$\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

5. Prove Hölder's inequality

$$\sum_{k=1}^{n} x_k y_k \leq \left(\sum_{k=1}^{n} x_k^p\right)^{1/p} \left(\sum_{k=1}^{n} y_k^q\right)^{1/q}, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

for $p > 1$, $x_k, y_k > 0$, by maximizing the function

$$\sum_{k=1}^{n} x_k y_k$$

in the x_k's for fixed, given y_k's, under the constraint

$$\sum_{k=1}^{n} x_k^p = b,$$

for a fixed, positive number b.

6. The Cobb-Douglas utility function u is of the form

$$u(x_1, x_2, x_3) = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}, \quad 0 \leq \alpha_i, \alpha_1 + \alpha_2 + \alpha_3 = 1, x_i \geq 0.$$

If a consumer has resources given by $\overline{x}_i$ for commodity X_i, and their respective prices are p_i, formulate and solve the problem of maximizing satisfaction measured through such utility function.

7. Find the best parabola through the four points $(-1, 4)$, $(0, 1)$, $(1, 1)$ and $(1, -1)$ by minimizing the least square error.

Many more examples could be stated. Readers should already know how to solve these problems.

After examining with some care all of these problems, we can draw some conclusions.

- Important ingredients of every optimization problem are the cost function, the function to be minimized or maximized, and the constraints that feasible vectors need to comply with to be even considered. The collection of these vectors is the feasible set of the problem.
- It is important, before starting to look for optimal solutions, to be sure that those are somewhere within the feasible set of the problem. This is the issue of the existence of optimal solutions. Some times optimization problems may not be well-posed, and may lack optimal solutions.
- In all of these examples, we deal with functions and vectors, no matter how many components they may have. The set of important techniques to treat this kind of optimization problems lie in the area universally known as Mathematical Programming and Operations Research.

The class of optimization problems we would like to consider in this text share some fundamental ingredients with those above: objective or cost functions, and constraints. But they differ fundamentally in the last feature: we would like to explore optimization problems for infinite dimensional vectors. Vectors are no longer vectors in the usual sense, but they become functions; and objective functions become functionals; and Mathematical Programming, as the discipline to deal with finite-dimensional optimization problems, becomes Calculus of Variations, one of the main areas of Analysis in charge of infinite dimensional optimization problems. Independent variables in problems, instead of being vectors, are, as already pointed out, functions. This passage from finite to infinite dimension is so big a change that it is not surprising that methods and techniques are so different from those at the heart of Mathematical Programming.

Let us start looking at some examples.

1.2 Basic Examples

Almost every formula to calculate some geometric or physical quantity through an integral can furnish an interesting problem in the Calculus of Variations. Our readers will know about the following typical cases:

1. The area enclosed by the graph of a non-negative function $u(x)$ between two values a and b:

$$A = \int_a^b u(x)\,dx.$$

2. The length of the graph of the function $u(x)$ between two values a and b:

$$L = \int_a^b \sqrt{1 + u'(x)^2}\, dx.$$

3. The volume of the solid of revolution generated by the graph of the function $u(x)$ around the X-axis between the two values a and b:

$$V = \pi \int_a^b u(x)^2\, dx.$$

4. The surface of revolution of the same piece of graph around the X-axis

$$S = 2\pi \int_a^b u(x)\sqrt{1 + u'(x)^2}\, dx.$$

5. The area of a piece of graph of a function $u(x, y)$ of two variables over a subset $\Omega \subset \mathbb{R}^2$ where the variables (x, y) move, is given by the integral

$$S = \int_\Omega \sqrt{1 + |\nabla u(x, y)|^2}\, dx\, dy. \tag{1.1}$$

Some examples coming from physical quantities follow.

1. The work done by a conservative force field $\mathbf{F}(\mathbf{x})$ where

$$\mathbf{F} = (F_1(\mathbf{x}), F_2(\mathbf{x}), F_3(\mathbf{x})), \quad \mathbf{x} = (x_1, x_2, x_3),$$

in going from a point $\mathbf{P}$ to another one $\mathbf{Q}$ is independent of the path followed, and corresponds to the difference of potential energy. But if the field is non-conservative, the work depends on the path, and is given by

$$W = \int_a^b \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\, dt, \quad \mathbf{P} = \mathbf{x}(a), \mathbf{Q} = \mathbf{x}(b).$$

2. The flux of a field $\mathbf{F}(\mathbf{x})$ through a surface $\mathbb{S}$ is given by

$$F = \int_{\mathbb{S}} \mathbf{F}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})\, dS(\mathbf{x}),$$

where $\mathbf{n}(\mathbf{x})$ is the normal to $\mathbb{S}$, and dS represents the element of surface area.
3. Moments of inertia. If a body of density $\rho(\mathbf{x})$ occupies a certain region Ω in space, given a certain axis $\mathbf{r}$, the moment of inertia with respect to $\mathbf{r}$, is given by

the integral

$$\int_{\Omega} r(\mathbf{x})\rho(\mathbf{x})\,d\mathbf{x},$$

where $r(\mathbf{x})$ is the distance from $\mathbf{x}$ to the axis $\mathbf{r}$.

Let us focus on one of these examples, say, the area enclosed by a portion of the graph of the function $u(x)$, $x \in (a, b)$, i.e.

$$A = \int_a^b u(x)\,dx. \tag{1.2}$$

There are essentially two ingredients in this formula: the two integration limits, and the function $u(x)$. We can therefore regard quantity A as a function of the upper limit b, maintaining the other ingredients fixed

$$A(b) = \int_a^b u(x)\,dx.$$

This is actually the viewpoint adopted when dealing with the Fundamental Theorem of Calculus to conclude that if $u(x)$ is continuous, then the "function" $A(b)$ is differentiable and $A'(b) = u(b)$. We can then think about the problem of finding the extreme values of $A(b)$ when b runs freely in $\mathbb{R}$, or in some preassigned interval $[\alpha, \beta]$.

But we can also think of A in (1.2) as a "function" of $u(x)$, keeping the two end-points a and b fixed, and compare the corresponding integrals for different choices of functions $u(x)$. If we clearly determine a collection $\mathcal{A}$ of such competing functions $u(x)$, we could ask ourselves which one of those functions realizes the minimum or the maximum of A. It is important to select in a sensible way the class $\mathcal{A}$ for otherwise the optimization may be pointless. Suppose we accept every function $u(x)$ one could think of, and try to figure out the one providing the maximum of the integral. It is pretty clear that there is no such function because the maximum is infinite: we can always find a function $u(x)$ for which the integral A is bigger than any preassigned number, no matter how big this is. This problem is useless. Assume, instead, that we only admit to compete for the minimum those functions $u(x)$ defined on the interval $[a, b]$, such that $u(a) = u(b) = 0$, and whose graph measures less than or equal to a positive number $L > b - a$. After a bit of experimentation, one realizes that the problem looks non-trivial and quite interesting, and that the function realizing this minimum, if there is one such function, must enjoy some interesting geometric property related to area and length of the graph.

We hope this short discussion may have helped in understanding the kind of situations that we would like to address in this text. We are interested in considering sets of functions; a way to assign a number to each one of those, a functional; decide

if there is one such function which realizes the minimum or maximum possible; and if there is one, find it, or derive interesting properties of it. This sort of optimization problems are identified as variational problems for reasons that will be understood later.

Almost as important as studying variational problems is to propose interesting, meaningful problems or cases. As we have seen with the above discussion, it is not always easy to make a proposal of a relevant variational problem. We are going to see next a bunch of those that have had a historical impact on Science and Engineering.

1.3 More Advanced Examples

The previous section has served to make us understand the kind of optimization problems we would like to examine in this text. We will describe next some of the paradigmatic examples that have played a special role historically, or that are among the distinguished set of examples that are used in most of the existing textbooks about the Calculus of Variations.

1. Transit problems.
2. Geodesics.
3. Dirichlet's principle.
4. Minimal surfaces.
5. Isoperimetric problems.
6. Hamiltonian mechanics.

We will see through these examples how important it is to be able to formulate a problem in precise mathematical terms that may enable to compare different alternatives, and decide which one is better; to argue if there are optimal solutions, and, eventually, find them. It is an initial, preliminary step that requires quite a good deal of practice, and which students usually find difficulties with.

1.3.1 Transit Problems

The very particular example that is universally accepted as marking the birth of the Calculus of Variations as a discipline on its own is the brachistochrone. We will talk about its relevance in the development of this field later in the final section of this chapter.

Given two points in a plane at different height, find the profile of the curve through which a unit mass under the action of gravity, and without slipping, employs the shortest time to go from the highest point to the lowest. Let $u(x)$ be one such

feasible profile so that

$$u(0) = 0, \quad u(a) = A,$$

with $a, A > 0$, and consider curves joining the two points $(0, 0)$, (a, A) in the plane. It is easy to realize that we can restrict attention to graphs of functions like the one represented by $u(x)$ because curves joining the two points which are not graphs cannot provide a minimum transit time under the given conditions. We need to express, for such a function $u(x)$, the time spent by the unit mass in going from the highest point to the lowest. From elementary kinematics, we know that

$$dt = \frac{ds}{\sqrt{2gh}}, \quad T = \int dt = \int \frac{ds}{\sqrt{2gh}},$$

where t is time, s is arc-length (distance), g is gravity, and h is height. We also know that

$$ds = \sqrt{1 + u'(x)^2}\, dx,$$

and can identify h with $u(x)$. Altogether we find that the total transit time T is given by the integral

$$T = \frac{1}{\sqrt{2g}} \int_0^a \frac{\sqrt{1 + u'(x)^2}}{\sqrt{u(x)}}\, dx.$$

An additional simplification can be implemented if we place the X-axis vertically, instead of horizontally, without changing the setup of the problem. Then we would have that

$$T = \frac{1}{\sqrt{2g}} \int_0^A \frac{\sqrt{1 + u'(x)^2}}{\sqrt{x}}\, dx, \tag{1.3}$$

and the problem we would like to solve is

$$\text{Minimize in } u(x): \quad \int_0^A \frac{\sqrt{1 + u'(x)^2}}{\sqrt{x}}\, dx$$

subject to

$$u(0) = 0, \quad u(A) = a.$$

Notice that this new function u is the inverse of the old one, and that positive, multiplicative constants do not interfere with the optimization process.

To stress the meaning of our problem, suppose, for the sake of definiteness, that we take $a = A = 1$. In this particular situation, we have three easy possibilities of a

function passing through the two points (0, 0), (1, 1), namely

$$u_1(x) = x, \quad u_2(x) = x^2, \quad u_3(x) = 1 - \sqrt{1 - x^2},$$

a straight line, a parabola, and an arc of a circle, respectively. Let us compare the transit time for the three, and decide which of the three yields a smaller value. According to (1.3), we have to compute (approximate) the value of the three integrals

$$T_1 = \int_0^1 \frac{\sqrt{2}}{\sqrt{x}}\,dx, \quad T_2 = \int_0^1 \frac{\sqrt{1+4x^2}}{\sqrt{x}}\,dx,$$
$$T_3 = \int_0^1 \frac{1}{\sqrt{x(1-x^2)}}\,dx.$$

Note that the three are improper integrals because the three integrands do have, at least, an asymptote at zero. Yet the value of the three integrals is finite. The smaller of the three is the one corresponding to the parabola u_2. But still the important issue is to find the best of all such admissible profiles.

1.3.2 Geodesics

It is well-known that geodesics in free, plane euclidean space are straight lines. However, when distances are distorted because of the action of some agent, then the closest curves joining two given points, the so-called geodesics, may not be the same straight lines. More specifically, if locally at a point $\mathbf{x} \in \mathbb{R}^N$ distances are measured by the formula

$$\|\mathbf{v}\|^2 = \mathbf{v}^T \mathbf{A}(\mathbf{x})\mathbf{v}$$

where $\mathbf{v} \in \mathbb{R}^N$, and the symmetric, positive-definite matrix $\mathbf{A}(\mathbf{x})$ changes with the point $\mathbf{x}$, then the total length of a curve

$$\mathbf{x}(t) : [0, 1] \to \mathbb{R}^N$$

would be given by the integral

$$\int_0^1 \left(\mathbf{x}'(t)^T \mathbf{A}(\mathbf{x}(t))\mathbf{x}'(t)\right)^{1/2}\,dt.$$

Hence, shortest paths between two points $\mathbf{P}_0$ and $\mathbf{P}_1$ would correspond to curves $\mathbf{X}$ realizing the minimum of such a problem

$$\text{Minimize in } \mathbf{x}(t): \quad \int_0^1 \left(\mathbf{x}'(t)^T \mathbf{A}(\mathbf{x}(t))\mathbf{x}'(t)\right)^{1/2} dt$$

subject to

$$\mathbf{x}(0) = \mathbf{P}_0, \quad \mathbf{x}(1) = \mathbf{P}_1.$$

The classical euclidean case corresponds to $\mathbf{A} = \mathbf{1}$, the identity matrix; but even for a situation in the plane $\mathbb{R}^2$ with

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

is not clear if geodesics would again be straight lines. What do you think?

1.3.3 Dirichlet's Principle

Another problem that played a major role in the development of the Calculus of Variations is the Dirichlet's principle in which we pretend to find the function

$$u(\mathbf{x}) : \Omega \subset \mathbb{R}^N \to \mathbb{R},$$

for a preassigned domain Ω, that minimizes the functional

$$\frac{1}{2} \int_\Omega |\nabla u(\mathbf{x})|^2 \, d\mathbf{x} \tag{1.4}$$

among all such functions $u(\mathbf{x})$ complying with $u = u_0$ around the boundary $\partial\Omega$ of Ω, and the function u_0 is given a priori. We are hence forcing functions to take some given values on around $\partial\Omega$. Note that $u(\mathbf{x})$ is a function of several variables, and

$$\nabla u(\mathbf{x}) = \left(\frac{\partial u}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial u}{\partial x_N}(\mathbf{x})\right), \quad |\nabla u(\mathbf{x})|^2 = \sum_{i=1}^N \frac{\partial u}{\partial x_i}(\mathbf{x})^2.$$

In certain circumstances, if graphs of such functions $u(\mathbf{x})$ are identified as shapes of elastic membranes fixed at the boundary $\partial\Omega$ according to the values of u_0, the integral in (1.4) provides a measure of the elastic energy repressed by the membrane when it adopts the shape determined by the graph of some $u(\mathbf{x})$. The minimizer would correspond to the function $u(\mathbf{x})$ whose graph would repressed the least energy

possible under such circumstances, and so it would yield the shape adopted by the membrane naturally.

1.3.4 Minimal Surfaces

One of the most fascinating variational examples is the one trying to minimize the area functional S in (1.1)

$$S(u) = \int_\Omega \sqrt{1 + |\nabla u(x, y)|^2}\, dx\, dy.$$

It is one of the main functionals that has stirred and is stirring a lot of research, both in Analysis and Geometry. Typically, the underlying variational problem tries to minimize the area $S(u)$ among those functions sharing the same values on the boundary $\partial\Omega$ of a given domain $\Omega \subset \mathbb{R}^2$. The graph of a given function $u(\mathbf{x})$ is a minimal-area surface if it is a minimizer of $S(u)$ among those functions sharing with u its boundary values around some Ω.

Because surface area is closely related to surface tension, minimal surfaces represent shapes adopted, for instance, by soap films as these are looking for surfaces minimizing surface tension.

1.3.5 Isoperimetric Problems

We know from Calculus, and has been reminded above, that many quantities associated with geometric objects can be expressed through integrals. The graph of a certain function $u(x)$ for $x \in [x_0, x_1]$ determines some important geometric quantities, except for positive multiplicative constants, like the area enclosed by it

$$A = \int_{x_0}^{x_1} u(x)\, dx;$$

the length of its graph

$$L = \int_{x_0}^{x_1} \sqrt{1 + u'(x)^2}\, dx;$$

the area of revolution around the X-axis

$$S = \int_{x_0}^{x_1} u(x)\sqrt{1 + u'(x)^2}\, dx;$$

the volume of the solid of revolution around the X-axis

$$V = \int_{x_0}^{x_1} u(x)^2 \, dx.$$

Typically, the optimization problem that consists in minimizing (or maximizing) one of these integrals under fixed end-point conditions is meaningless either because the optimal answer can be easily found, and the problem becomes trivial; or because the corresponding extreme value is not finite. For instance, finding the minimum of the length of a graph passing through two given points in the plane is trivially found to be the straight line through those two points, while the maximum length is infinite.

It is a much more interesting situation to use two of those functionals to setup interesting and important variational problems. The classical one is the following:

Find the optimal function $u(x)$ for $x \in [-1, 1]$ that minimizes the integral

$$\int_{-1}^{1} u(x) \, dx$$

among those with

$$u(-1) = u(1) = 0,$$

and having a given length $L > 2$ of its graph, i.e. among those respecting the condition

$$L = \int_{-1}^{1} \sqrt{1 + u'(x)^2} \, dx.$$

This is a much more interesting and fascinating problem. In general terms, an isoperimetric problem is one in which we try to minimize an integral functional among a set of functions which are restricted by demanding, among other possible constraints, a condition setup with another integral functional. There are various possibilities playing, for example, with the quantities A, L, S and V, above.

There are, in particular, two very classical situations.

1. Dido's problem. Given a rope of length $L > 0$, determine the maximum area that can be enclosed by it, and the shape to do so. It admits some interesting variants. The following is an elementary way to state the situation. Let $(u(t), v(t))$ be the two components of a plane, closed curve σ parameterized in such a way that

$$u(a) = u(b) = 0, \quad v(a) = v(b) = 0.$$

The length of σ is given by

$$L(\sigma) = \int_a^b \sqrt{u'(t)^2 + v'(t)^2} \, dt,$$

while the area enclosed by it is, according to Green's theorem,

$$A(\sigma) = \int_a^b (u(t)v'(t) - u'(t)v(t))\,dt.$$

We would like to find the optimal curve for the problem

$$\text{Minimize in } \sigma: \quad L(\sigma) \quad \text{under } A(\sigma) = \alpha, \sigma(a) = \sigma(b) = (0, 0).$$

2. The hanging cable. This time we have a uniform, homogeneous cable of total length L that is to be suspended from its two end-points between two points at the same height, and separated a distance H apart. We will necessarily have $L > H$. We would like to figure out the shape that the hanging cable will adopt under the action of its own weight. If we assume that such a profile will be the result of minimizing the potential energy associated with any such admissible profile represented by the graph of a function $u(x)$, then we know that potential energy is propotional to height, and so

$$dP = u(x)\,ds, \quad ds = \sqrt{1 + u'(x)^2}\,dx.$$

The full potential energy contained in such feasible profile will then be

$$P = \int_0^H u(x)\sqrt{1 + u'(x)^2}\,dx.$$

Constraints now should acount for the fact that the cable has total length L, in addition to demanding

$$u(0) = 0, \quad u(H) = 0.$$

The constraint on the length reads

$$L = \int_0^H \sqrt{1 + u'(x)^2}\,dx,$$

coming from integrating in the full interval $[0, H]$ arc-length ds. We then seek to find the optimal shape corresponding to the problem

$$\text{Minimize in } u(x): \quad \int_0^H u(x)\sqrt{1 + u'(u)^2}\,dx$$

under the constraints

$$u(0) = u(H) = 0, \quad L = \int_0^H \sqrt{1 + u'(x)^2}\,dx.$$

1.3.6 Hamiltonian Mechanics

There is a rich tradition of variational methods in Mechanics. If $\mathbf{x}(t)$ represents the state of a certain mechanical system with velocities $\mathbf{x}'(t)$, then the dynamics of the system evolves in such a way that the action integral

$$\int_0^T L(\mathbf{x}(t), \mathbf{x}'(t))\, dt$$

is minimized. The integrand

$$L(\mathbf{x}, \mathbf{y}) : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$$

is called the lagrangian of the system. The hamiltonian $H(\mathbf{x}, \mathbf{y})$ is defined through the formula

$$H(\mathbf{x}, \mathbf{y}) = \sup_{\mathbf{z}\in\mathbb{R}^N} \{\mathbf{z} \cdot \mathbf{y} - L(\mathbf{x}, \mathbf{y})\}.$$

H, defined through this formula, is called the conjugate function of L with respect to the variable $\mathbf{y}$, as variable $\mathbf{x}$ here plays the role of a vector of parameters. It is interesting to note that, under hypotheses that we do not bother to specify,

$$L(\mathbf{x}, \mathbf{y}) = \sup_{\mathbf{z}\in\mathbb{R}^N} \{\mathbf{z} \cdot \mathbf{y} - H(\mathbf{x}, \mathbf{y})\}$$

as well.

We will learn later to write the Euler-Lagrange (E-L) system that critical paths of the functional with integrand $L(\mathbf{x}, \mathbf{y})$ above ought to verify. It reads

$$-\frac{d}{dt} L_{\mathbf{y}}(\mathbf{x}(t), \mathbf{x}'(t)) + L_{\mathbf{x}}(\mathbf{x}(t), \mathbf{x}'(t)) = \mathbf{0}. \tag{1.5}$$

Paths which are solutions of this system will also evolve according to the hamiltonian equations

$$\mathbf{y}'(t) = -\frac{\partial H}{\partial \mathbf{x}}(\mathbf{x}(t), \mathbf{y}(t)), \quad \mathbf{x}'(t) = \frac{\partial H}{\partial \mathbf{y}}(\mathbf{x}(t), \mathbf{y}(t)).$$

In fact, the definition of the hamiltonian leads to the relations

$$H(\mathbf{x}, \mathbf{y}) = \mathbf{z}(\mathbf{x}, \mathbf{y}) \cdot \mathbf{y} - L(\mathbf{x}, \mathbf{z}(\mathbf{x}, \mathbf{y})), \quad \mathbf{y} - L_{\mathbf{y}}(\mathbf{x}, \mathbf{z}(\mathbf{x}, \mathbf{y})) = \mathbf{0},$$

where this $\mathbf{z}$ is the optimal vector $\mathbf{z}$ (depending on $\mathbf{x}$ and $\mathbf{y}$) in the definition of the hamiltonian. Then it is easy to check that

$$L_{\mathbf{y}}(\mathbf{x}, \mathbf{z}) = \mathbf{y}, \quad \mathbf{z} = H_{\mathbf{y}}(\mathbf{x}, \mathbf{y}), \quad H_{\mathbf{x}}(\mathbf{x}, \mathbf{y}) = -L_{\mathbf{x}}(\mathbf{x}, \mathbf{z}),$$

If we put

$$\mathbf{x}' = \mathbf{z}, \quad \mathbf{y} = L_{\mathbf{y}}(\mathbf{x}, \mathbf{x}'),$$

it is immediate to check, through these identities, the equivalence between the E-L system and Hamilton's equations. Note that system (1.5) is

$$\mathbf{y}' = -H_{\mathbf{x}}(\mathbf{x}, \mathbf{y}).$$

1.4 The Model Problem, and Some Variants

We have emphasized in the preceding sections the interest, both from a purely analytical but also from a more applied standpoint, of studying variational problems of the general form

$$I(\mathbf{u}) = \int_{\Omega} F(\mathbf{x}, \mathbf{u}(\mathbf{x}), \nabla \mathbf{u}(\mathbf{x}))\, d\mathbf{x}$$

under typical additional conditions like limiting the value of competing fields $\mathbf{u}(\mathbf{x})$ around $\partial\Omega$. Here we have:

- $\Omega \subset \mathbb{R}^N$ is a typical subset, which cannot be too weird;
- feasible fields

  $$\mathbf{u}(\mathbf{x}) : \Omega \to \mathbb{R}^n$$

  should be differentiable with gradient or differential (jacobian) matrix

  $$\nabla \mathbf{u}(\mathbf{x}) : \Omega \to \mathbb{R}^{n \times N};$$

- the integrand

  $$F(\mathbf{x}, \mathbf{u}, \mathbf{z}) : \Omega \times \mathbb{R}^N \times \mathbb{R}^{n \times N} \to \mathbb{R}$$

  is assumed to enjoy smoothness properties that will be specified as they are needed.

Dimensions N and n make the problem quite different ranging from understandable to pretty hard. We will refer to:

1. $N = n = 1$: scalar ($n = 1$), variational problems in dimension $N = 1$;
2. $N = 1, n > 1$: vector problems in dimension one;
3. $N > 1, n = 1$: scalar, multidimensional problems;
4. $N, n > 1$: vector, multidimensional problems.

In general terms we will refer to the scalar case when either of the two dimensions N or n is 1, while we will simply classify a problem as vectorial if both dimensions are greater than unity. We will focus mainly in the scalar case in this text, while the vector case in left for another book. Note how all of the examples examined earlier in the chapter fall into the category of scalar problems.

To stress how variational problems of the above class are just a first step to more general situations of indisputable interest in Science and Engineering, consider a scalar, uni-dimensional variational problem

$$\text{Minimize in } u(t): \quad \int_0^T F(t, u(t), u'(t))\,dt$$

under the end-point conditions

$$u(0) = u_0, \quad u(T) = u_T.$$

It is obvious that we can also write the problem in the form

$$\text{Minimize in } v(t): \quad \int_0^T F(t, u(t), v(t))\,dt$$

under the constraints

$$u(0) = u_0, \quad u(T) = u_T, \quad u'(t) = v(t).$$

Notice how we are regarding the function $v(t)$ as our "variable" for the optimization problem, while $u(t)$ would be obtained from $v(t)$ through integration. In particular, there is the integral constraint

$$\int_0^T v(t)\,dt = u_T - u_0$$

that must be respected by feasible functions $v(t)$. The obvious relationship

$$u'(t) = v(t)$$

can immediately be generalized to

$$u'(t) = f(t, u(t), v(t))$$

for an arbitrary function

$$f(t, u, v) : (0, T) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}.$$

This is not just a more general situation for the sake of it, since it gives rise to a new class of optimization problems of a tremendous impact in Engineering that are identified as optimal control problems (of ODEs).

There is another important generalization that our readers should keep in mind. So far integrands in functionals have been assumed to depend explicitly on the first derivative u' or gradient ∇u of competing functions. This does not have to be so. In fact, the highest derivative occurring explicitly in a variational problem indicates the order of the problem. Usually, first-order problems are the ones studied because these families of problems are the most common. However, one can consider zero-order, and higher-order problems. A functional of the form

$$\int_0^1 F(t, u(t))\, dt$$

would correspond to a zero-order problem, assuming that no derivative participates explicitly in additional constrains, whereas

$$\int_0^1 F(t, u(t), u'(t), u''(t))\, dt$$

with a explicit dependence of F on the variable u'' would indicate a second-order example.

1.5 The Fundamental Issues for a Variational Problem

What are the main concerns when facing any of the examples we have just examined, or any other such problem for that matter? The first and fundamental issue is to know if there is a particular feasible function realizing the minimum. If there is such a function (there might be several of them), it will be quite special in some sense. How can one go about showing whether there is a minimizer for one of our variational problems? Let us just backup to a finite dimensional situation to see what we can learn from there.

Suppose we have a function

$$F(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R},$$

and we would like to convince ourselves that there is at least one vector $\mathbf{x}_0 \in \mathbb{R}^N$ with the property

$$F(\mathbf{x}_0) = \min_{\mathbf{x}\in\mathbb{R}^N} F(\mathbf{x}).$$

We would probably look immediately, specially if we were good Calculus students, at the system of critical points

$$\nabla F(\mathbf{x}) = \mathbf{0}. \tag{1.6}$$

That is not a bad idea. However, we are well aware that, even if $\mathbf{x}_0$ must be a solution of this non-linear system, provided F is differentiable, there might be other solutions for (1.6). On the other hand, we do not have the slightest idea about what the equivalent to the critical point system (1.6) would be for a functional like the one in the previous section. In case F is not differentiable, then we would feel at a loss.

There are some precautions one should always take before moving on. Let us put

$$m = \inf_{\mathbf{x}\in\mathbb{R}^N} F(\mathbf{x}).$$

We would definitely like $m \in \mathbb{R}$ to be a real number. It could however happen that $m = -\infty$ (the case $m = +\infty$ does not make much sense, why?). If $m = -\infty$, it means that the problem is not well-posed in the sense that the values of F can decrease indefinitely, and moreover, if F is continuous,

$$\liminf_{|\mathbf{x}|\to\infty} F(\mathbf{x}) = -\infty.$$

These situations are not interesting as we could not have minimizers. How can one avoid them? A standard condition is known as coercivity: such a function F is said to be coercive if, on the contrary,

$$\liminf_{|\mathbf{x}|\to\infty} F(\mathbf{x}) = +\infty. \tag{1.7}$$

This condition suffices to ensure that $m \in \mathbb{R}$.

Once we know that $m \in \mathbb{R}$, one thing one can always find (establish the existence of) is a minimizing sequence $\{\mathbf{x}_j\}$ for F. This is simply a sequence of vectors $\{\mathbf{x}_j\}$ such that

$$F(\mathbf{x}_j) \searrow m = \inf_{\mathbf{x}\in\mathbb{R}^N} F(\mathbf{x}).$$

This is nothing but the definition itself of the infimum of a collection of numbers. Another matter is how to find in practice, for specific functions F, such a sequence.

In some sense that minimizing sequence should be telling us where to look for minimizers. What is sure, under the coerciveness condition (1.7), is that $\{\mathbf{x}_j\}$ is a bounded, collection of vectors. As such, it admits some converging subsequence (Heine-Borel theorem) $\{\mathbf{x}_{j_k}\}$:

$$\mathbf{x}_{j_k} \to \mathbf{x}_0 \text{ as } k \to \infty.$$

This is a fundamental compactness property. If F is continuous, then

$$F(\mathbf{x}_0) = \lim_{k\to\infty} F(\mathbf{x}_{j_k}) \searrow m,$$

and $\mathbf{x}_0$ becomes one desired minimizer. Once we are sure that there are vectors where the function F achieves its minimum value, then we know that those must be solutions for the system of critical points (1.6), and we can start from here to look for those minimum points.

Let us mimic this process with a typical integral functional of the form

$$I(u) = \int_0^1 F(u(x), u'(x))\, dx \tag{1.8}$$

where

$$F(x, z) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad u(x) : [0, 1] \to \mathbb{R},$$

and there could be further conditions, most probably in the form of end-point restrictions, for admissible functions. The density F is assumed to be, to begin with, just continuous, and for the sake of definiteness we have taken the interval of integration to be the unit interval $[0, 1] \subset \mathbb{R}$. The functional I is supposed not to be identically $+\infty$, i.e. there is a least one admissible function u with $I(u) \in \mathbb{R}$. Set

$$m = \inf_{\mathcal{A}} I(u),$$

where $\mathcal{A}$ stands for the class of functions that we would allow in the minimization process, essentially C^1-functions (in order to have a derivative) complying with additional restrictions. The first concern is to guarantee that $m \in \mathbb{R}$, i.e. $m > -\infty$. In the case of a function of several variables $F(\mathbf{x})$, we discussed above that the coercivity condition (1.7) suffices to ensure this property. What would the substitute be for a functional like $I(u)$? Note how this requires to determine in a precise way what it means for a function

$$u(x) : [0, 1] \to \mathbb{R}$$

to go to "infinity". In the case of a finite-dimensional vector $\mathbf{x}$, it clearly means that

$$|\mathbf{x}|^2 = \sum_{i=1}^{N} x_i^2 \to \infty, \quad \mathbf{x} \in \mathbb{R}^N, \mathbf{x} = (x_1, x_2, \ldots, x_N).$$

One can think of the condition, which is possibly the most straightforward generalization of this finite-dimensional case,

$$\sup_{x\in[0,1]} |u(x)| \to \infty.$$

But there might be other possibilities like demanding, for instance,

$$\int_0^1 |u(x)|\,dx \to \infty.$$

What is pretty clear is that a certain "measure" of the overall size of the function must tend to infinity, and as the size of a feasible function $u \in \mathcal{A}$ tends to infinity, the numbers $I(u)$ ought to go to infinity as well.

Assume that we have somehow solved this problem, determined clearly how to calculate the global size $|u| \in \mathbb{R}^+$ of a function $u \in \mathcal{A}$, and checked that the functional $I(u)$ is coercive in the sense

$$\liminf_{|u|\to\infty} I(u) = +\infty,$$

in such a way that minimizing sequences $\{u_j\} \subset \mathcal{A}$ enjoy the compactness property

$$\{|u_j|\} \subset \mathbb{R},$$

a bounded collection of numbers. At this stage, in the finite-dimensional situation, one resorts to the fundamental compactness property that there should be subsequences converging to some vector. This innocent-looking assertion encloses various deep concepts. We need to grasp how to find a replacement of this compactness property in the infinite-dimensional scenario.

Given that functions $u(x)$ are in fact numbers for each point in the domain $x \in [0, 1]$, the most direct way to argue is possibly as follows. If the property $\{|u_j|\} \subset \mathbb{R}$ implies that $\{u_j(x)\}$ is a bounded sequence of numbers for every such x, we can certainly apply the compactness property in $\mathbb{R}$ to conclude that there will be a certain subsequence j_k such that $u_{j_k}(x) \to u(x)$ where we have put $u(x)$ for the limit, anticipating thus the limit function. There are two crucial issues:

1. The subsequence j_k which is valid for a certain $x_1 \in [0, 1]$ might not be valid for another point $x_2 \in [0, 1]$, that is to say, the choice of j_k may depend on the point x. We can start taking subsequences of subsequences in a typical diagonal argument, but given that the interval $[0, 1]$ is not countable, there is no

way that this process could lead anywhere. Unless we find a way, or additional requirements in the situation, to tune the subsequence j_k for all points $x \in [0, 1]$ at once so that

$$u_{j_k}(x) \to u(x) \text{ for all } x \in [0, 1],$$

it is impossible to find a limit function in this way.

2. Even if we could resolve the previous step, there is still the issue to check whether the limit function $u(x) \in \mathcal{A}$. In particular, u must be a function for which the functional I can be computed.

Example 1.1 To realize that these difficulties are not something fancy or pathological, let us examine the following elementary example. Put

$$u_j(x) = \sin(2^j \pi x), \quad x \in [0, 1].$$

It is not difficult to design variational problem like (1.8) for which sequences similar to u_j are minimizing. For each $x \in [0, 1]$, $\{u_j(x)\}$ is obviously a sequence of number in the interval $[-1, 1]$, and so bounded. But there is no way to tune a single subsequence j_k to have that $\{\sin(2^{j_k} \pi x)\}$ is convergent for all x at once. Note that the set of numbers

$$\{\sin(2^j \pi x) : j \in \mathbb{N}\}$$

is dense in $[-1, 1]$ for every $x \in (0, 1)$. In this case, there does not seem to be a reasonable way to define a limit function for this sequence. The solution to this apparent dead-end is surprising and profound.

To summarize our troubles, we need to address the following three main issues:

1. we need to find a coherent way to define the size of a function; there might be various ways to do so;
2. we need some compactness principle capable of determining a function from a sequence of functions whose sizes form a bounded sequence of positive numbers;
3. limit functions of bounded-in-size sequences should enjoy similar properties to the members of the sequences.

Before we can even talk about how to tackle variational problems like (1.8), there is no way out but to address these three fundamental issues. This is the core of our initial exposure to Functional Analysis.

1.6 Additional Reasons to Care About Classes of Functions

Let us start making some explicit calculations to gain some familiarity with simple variational principles.

Let us focus on some scalar, one-dimensional examples of the form

$$I(u) = \int_{x_0}^{x_1} F(x, u(x), u'(x))\, dx. \tag{1.9}$$

Note how in this case

$$\Omega = (x_0, x_1), \quad x \in (x_0, x_1),$$
$$u(x) : (x_0, x_1) \to \mathbb{R}, \quad u'(x) : (x_0, x_1) \to \mathbb{R},$$

and additional standard conditions will most likely limit the values of u or its derivative at one or both end-points

$$x = x_0, \quad x = x_1.$$

Different problems will obviously depend on the integrand

$$F(x, u, z) : (x_0, x_1) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}.$$

To compute an integral like the one in (1.9), for a given function $u(x)$, we clearly need to ensure that u is differentiable in (x_0, x_1), and plug its derivative $z = u'(x)$ as the third argument in F; hence, we must restrict ourselves to look for the infimum of I in (1.9) among differentiable functions.

Let us explore the following two examples:

1. the variational problem

$$\text{Minimize in } u(x) : \quad \int_0^1 u(x)^2 (2 - u'(x))^2\, dx$$

   subjected to

$$u(0) = 0, \quad u(1) = 1;$$

2. the variational problem

$$\text{Minimize in } u(x) : \quad \int_{-1}^1 u'(x)^6 (u(x)^3 - x)^2\, dx$$

   subjected to

$$u(-1) = -1, \quad u(1) = 1.$$

Concerning the first, we notice that the integrand is non-negative, being a square, and so it would vanish for functions $u(x)$ so that

$$u(x)(2 - u'(x)) = 0$$

for every point $x \in (0, 1)$. Hence, either $u(x) = 0$, or else $u'(x) = 2$. Given the end-point conditions

$$u(0) = 0, \quad u(1) = 1,$$

that we must respect, it is easy to conclude that

$$\overline{u}(x) = \begin{cases} 0, & 0 \le x \le 1/2, \\ 2x - 1, & 1/2 \le x \le 1, \end{cases}$$

is one possible optimal solution of the problem. But $\overline{u}$ is not differentiable !! The same problem pops up for any minimizer. If we insist in restricting attention to differentiable functions, $m = 0$ is the unattainable infimum of the problem: $\overline{u}$ can be approximated by C^1-functions in an arbitrary way by rounding the corner of the graph of $\overline{u}$, but the value $m = 0$ can never be reached. The situation is like solving the equation $x^2 = 2$ in the field of rational numbers.

The second example is more sophisticated. It is a typical situation of the phenomenon discovered by Lavrentiev in the 1920s. As in the first example, the integrand is non-negative, and the only way to lead integrals to zero, under the given end-point conditions, is by taking

$$\overline{u}(x) = \sqrt[3]{x}. \tag{1.10}$$

This unique minimizer is not differentiable either at $x = 0$. However, the situation is more dramatic than in the previous example. Suppose we pretend to approximate the minimizer in (1.10) by a family of differentiable functions. The most straightforward way would be to select a small $h > 0$, and use a piece of a straight line interpolating the two points $(-h, \sqrt[3]{-h})$ and $(h, \sqrt[3]{h})$ in the interval $[-h, h]$. The slope of this secant is

$$\frac{\sqrt[3]{h} - \sqrt[3]{-h}}{2h} = \frac{1}{\sqrt[3]{h^2}}.$$

The resulting function is not differentiable either at $\pm h$, but those two discontinuities can be rounded off without any trouble. The most surprising fact is that if we put

$$I(u) = \int_{-1}^{1} u'(x)^6 (u(x)^3 - x)^2 \, dx$$

and

$$u_h(x) = \begin{cases} \overline{u}(x), & -1 \le x \le -h, h \le x \le 1, \\ h^{-2/3}x, & -h \le x \le h, \end{cases},$$

then, due to evenness,

$$I(u_h) = 2\int_0^h h^{-4}(h^{-2}x^3 - x)^2\,dx.$$

A careful computation yields

$$I(u_h) = \frac{2}{h}\left(\frac{1}{3} + \frac{1}{7} - \frac{2}{5}\right),$$

and we clearly see that $I(u_h) \to \infty$ as $h \to 0$!! It is a bit shocking, but this elementary computation indicates that the behavior of this apparently innocent-looking functional is quite different depending on the set of competing functions considered. In particular, and this would require some more work than this simple computation we have just performed, the value of the infimum of $I(u)$ over a class of piece-wise affine functions over $[-1, 1]$ cannot vanish. And yet, when we permit the function $\overline{u}(x)$ to compete for the infimum, the minimum vanishes !!

What we have tried to emphasize with these two examples is the fundamental relevance in variational problems to specify the class of functions in which we are seeking to minimize a given functional. Historically, it took quite a while to realize this, but variational problems were at the core to push mathematicians to formalize spaces of functions where "points" are no longer vectors of a given dimension, but rather full functions. This led to the birth of an independent, fundamental discipline in Mathematics known universally as Functional Analysis.

1.7 Finite Versus Infinite Dimension

One of the main points that students must make an effort in understanding is the crucial distinction between finite and infinite-dimension. Though this is not always an issue, Functional Analysis focuses on discovering the places where an infinite-dimensional scenario is different, even drastically different, and what those main differences are or might be. To master where caution is necessary takes time because genuine infinite-dimensional events may be some times counter-intuitive, and one needs to educate intuition in this regard.

To stress this point, let us present two specific examples. We will see that the first one is identical in finite and infinite dimension, whereas in the second, the framework of infinite dimension permits new phenomena that need to be tamed.

This is part of our objective with this text. For the next statement, it is not indispensable to know much about metric spaces as the proof is pretty transparent. We will recall the basic definitions in the next chapter.

Theorem 1.1 (The Contraction Mapping Principle) *Let $\mathbb{H}$ be a complete metric space under the distance function*

$$d(\mathbf{x}, \mathbf{y}) : \mathbb{H} \times \mathbb{H} \to \mathbb{R}^{+}.$$

Suppose the mapping $\mathbf{T} : \mathbb{H} \to \mathbb{H}$ is a contraction in the sense

$$d(\mathbf{Tx}, \mathbf{Ty}) \leq K d(\mathbf{x}, \mathbf{y}), \quad 0 < K < 1. \tag{1.11}$$

There there is a unique fixed point for $\mathbf{T}$, i.e. a unique $\overline{\mathbf{x}} \in \mathbb{H}$ such that $\mathbf{T}\overline{\mathbf{x}} = \overline{\mathbf{x}}$.

Proof Select, in an arbitrary way, $\mathbf{x}_0 \in \mathbb{H}$, and define recursively the sequence

$$\mathbf{x}_k = \mathbf{T}\mathbf{x}_{k-1}, \quad k \geq 1.$$

We claim that $\{\mathbf{x}_k\}$ is a Cauchy sequence. In fact, if $j < k$, by the triangle inequality and the repeated use of the contraction inequality (1.11),

$$\begin{aligned} d(\mathbf{x}_k, \mathbf{x}_j) \leq & \sum_{l=j}^{k-1} d(\mathbf{x}_l, \mathbf{x}_{l-1}) \\ \leq & \sum_{l=j}^{k-1} K^l d(\mathbf{T}\mathbf{x}_0, \mathbf{x}_0) \\ = & d(\mathbf{T}\mathbf{x}_0, \mathbf{x}_0) \sum_{l=j}^{k-1} K^l \\ = & d(\mathbf{T}\mathbf{x}_0, \mathbf{x}_0) \frac{K^j - K^k}{1 - K}. \end{aligned}$$

Since $K < 1$, we clearly see that $d(\mathbf{x}_k, \mathbf{x}_j) \to 0$ as $k, j \to \infty$, and thus, the sequence converges to a limit element $\overline{\mathbf{x}}$. Since

$$d(\mathbf{T}^{(k+1)}\mathbf{x}_0, \mathbf{T}\overline{\mathbf{x}}) \leq K d(\mathbf{T}^{(k)}\mathbf{x}_0, \overline{\mathbf{x}}),$$

the left-hand side converges to $d(\overline{\mathbf{x}}, \mathbf{T}\overline{\mathbf{x}})$, and the right-hand side converges to 0, we conclude that indeed $\overline{\mathbf{x}}$ is a fixed point of $\mathbf{T}$. The same inequality (1.11) exclude the possibility of having more than one such fixed point.

Note how the statement and the proof of such a result does not make the distinction between finite and infinite dimension. As a general rule, when a fact can be shown in

a metric space regardless of any other underlying structure, that result will be valid regardless of dimension.

Possibly, from a more practical viewpoint, one of the main differences between finite and infinite-dimension refers to compactness. This is always a headache in Applied Analysis, and the main reason for the introduction and analysis of weak convergence. In some sense, in an infinite dimension context there are far too many dimensions where things may hide away or vanish. We refer to Example 1.1. For the sequence of trigonometric functions

$$\{\sin(jx)\}, \quad x \in [-\pi, \pi],$$

there are at least two facts that strikes when compared to a situation in finite dimension:

1. there cannot be a subsequence converging in a natural manner to anything;
2. any finite subset (of an arbitrary number of elements) of those sin-functions is always linearly independent.

We have already emphasized the first fact earlier. Concerning the second, suppose we could find that

$$\sum_{j=1}^{k} \lambda_j \sin(jx) = 0, \quad x \in [-\pi, \pi]$$

for a collection of numbers λ_j, and arbitrary k. If we multiply this identity by $\sin(lx)$, $1 \le l \le k$, and integrate in $[-\pi, \pi]$, we would definitely have

$$0 = \sum_{j=1}^{k} \lambda_j \int_{-\pi}^{\pi} \sin(jx)\sin(lx)\,dx.$$

But an elementary computation yields that

$$\int_{-\pi}^{\pi} \sin(jx)\sin(lx)\,dx = 0$$

when $j \neq l$. Hence

$$0 = \lambda_l \int_{-\pi}^{\pi} \sin^2(lx)\,dx,$$

which implies $\lambda_l = 0$ for all l. This clearly means that such an infinite collection of trigonometric functions is independent: the dimension of any vector space containing them must be infinite-dimensional.

1.8 Brief Historical Background

The development and historical interplay between Functional Analysis and the Calculus of Variations is one of the most fascinating chapters of the History of Mathematics. It looks, then, appropriate to write a few paragraphs on this subject addressed to young students, with the sole purpose to ignite the spark of curiosity. It may be a rather presumptuous attitude on our part, twenty-first century mathematicians, to look back on History and marvel at the difficulties that our former colleagues found in understanding concepts and ideas that are so easily conveyed today in advanced mathematics lectures all over the world. We should never forget, though, that this easiness in grasping and learning deep and profound concepts and results is the outcome of a lot of work and dedication by many of the most brilliant minds of the XIX- and XX-centuries. Modesty, humility, and awe must be our right attitudes. Our comments in this section are essentially taken from [BiKr84, Kr94I, Kr94II]. In [FrGi16], there is a much more detailed description of the early history of the Calculus of Variations, and its interplay with Functional Analysis and other areas. See also [Go80].

The truth is that concepts that are so much ingrained in our mathematics mentality today like that of "function" took quite a while until it was recognized and universally adopted. Great mathematicians like Euler, Lagrange, Fourier, Dirichlet, Cauchy, among others, but especially Riemann and Weierstrass contributed in a very fundamental way. The concept of "space" was, however, well behind that of function. The remarkable rise of non-euclidean geometries (Gauss, Lobachevsky, Bolyai, Klein) had a profound impact on the idea of what might be a "space". Some help came from Mechanics through the work of Lagrange, Liouville, Hamilton, Jacobi, Poincaré, …; and from Geometry (Cayley, Grassmann, Riemann, etc). It was Riemann who actually introduced, for the first time in his doctoral thesis in 1851, the idea of a "function space". There were other pioneers like Dedekind and Méray, but it was the fundamental work of Cantor with his set theory that permits the whole building of Functional Analysis as we think about this subject today. We cannot forget the italian school as it is considered the place where Functional Analysis started at the end of nineteenth century, and the initial part of twentieth century. Names like Ascoli, Arzelá, Betti, Beltrami, Cremona, Dini, Peano, Pincherle, Volterra, should not be forgotten. But also the french school, in the heart of the twentieth century, contributed immensely to the establishment of Functional Analysis as a discipline on its own. Here we ought to bring to mind Baire, Borel, Darboux, Fréchet, Goursat, Hadamard, Hermite, Jordan, Lebesgue, Picard, among many others.

The Calculus of Variations, and the theory of integral equations for that matter, provided the appropriate ground for the success of functional analytical techniques in solving quite impressive problems. Variational problems, formulated in a somewhat primitive form, started almost simultaneously with the development of Calculus. Yet the issue of making sure that there was a special curve or surface furnishing the minimum value possible for a certain quantity (a functional defined

on a space) was out of the question until the second half of the nineteenth century. It was not at all easy to appreciate the distinction between the unquestionable physical evidence of the existence of minimizers in Mechanics, and the need for a rigorous proof that it is so.

There are two fundamental years for the Calculus of Variations: 1696 is universally accepted as the birth of the discipline with J. Bernoulli's brachistochrone problem; while 1744 is regarded as the beginning of its theory with the publication by Euler of his necessary condition for a minimum. It was Lagrange who pushed Euler's ideas beyond, and invented the "method of variations" that gave its name to the discipline. Later Jacobi and Weierstrass discovered the sufficiency conditions for an extreme, and Legendre introduced the second variation.

One of the most fundamental chapters of the mutual interplay between the Calculus of Variations and Functional Analysis was Dirichlet's principle, or more generally Partial Differential Equations. It reads:

There exists a function $u(\boldsymbol{x})$ that minimizes the funcional

$$D(u) = \int_\Omega |\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x},$$

for $\Omega \subset \mathbb{R}^N$, among all continuously differentiable functions in Ω, and continuous up to the boundary $\partial\Omega$, taking prescribed values around $\partial\Omega$ given by f. Such function $u(\boldsymbol{x})$ is the solution of the problem

$$\Delta u = 0 \text{ in } \Omega, \quad u = f \text{ on } \partial\Omega, \quad u \in C^2(\Omega).$$

Its rigorous analysis had an unquestionable impact on Analysis, and it led to several interesting methods to show the existence of such a function minimizer u by Neumann, Poincaré, and Schwarz. But it also stirred ideas, initiated by Hilbert, that culminated in the direct method of the Calculus of Variations. These methods do not deal with the Euler-Lagrange equation or system, but place themselves in a true functional analytical setting by "suitably generalizing the concept of solution". This idea, expressed by Hilbert in a paper in 1900, became a guide for the Calculus of Variations for the twentieth century, and culminated in an important first step, along with the notion of semicontinuity, in the work of Tonelli at the beginning of the last century. There were other fundamental problems that also had a very clear influence in Analysis: Plateau's problem and minimal surfaces, and Morse theory. These areas started to attract so much attention that they eventually developped to be independent on their own. They are part of a different story, just as it is so with the field of integral equations.

It was precisely the important work of Fredholm on integral equations that attracted the interest of Hilbert and other colleagues in Göttingen. A lot of work was displayed on integral equations on those first decades of the twentieth century in Göttingen where a large number of Hilbert's disciples, comprising the Hilbert school, simplified, further explored and even extended Hilbert's intuition, and paved the way towards the year 1933 where Functional Analysis was started to being

clearly considered an independent, important branch of Mathematics. There are some other personal, fundamental contributions to Functional Analysis that cannot be justly missed or ignored: Banach, Hahn, and Riesz; and some later figures like Schauder, Stone, or Von Neumann.

This is just a glimpse on the origins of these two fundamental areas of Analysis. We have not touched on ideas, since that would require a full treatise. We have not mentioned every mathematician deserving it. Again brevity has forced us to name just the characters regarded as key figures in the early development of these parts of Mathematics, hoping to have motivated on readers a genuine interest in learning more.

1.9 Exercises

1. Prove Hölder's inequality

$$\sum_{k=1}^{n} x_k y_k \le \left(\sum_{k=1}^{n} x_k^p\right)^{1/p} \left(\sum_{k=1}^{n} y_k^q\right)^{1/q}, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

for $p > 1$, $x_k, y_k > 0$.

(a) Argue that the maximization problem

$$\text{Maximize in } \mathbf{x} = (x_1, x_2, \dots, x_n) : \quad \mathbf{x} \cdot \mathbf{y}$$

subjected to the condition

$$\sum_{k=1}^{n} x_k^p = b$$

where

$$b > 0, \quad \mathbf{y} = (y_1, y_2, \dots, y_n),$$

are given, has a unique solution.

(b) Find such solution.

(c) By interpreting appropriately the maximal property of the found solution, show Hölder's inequality.

2. For the pth-norm ($p \ge 1$)

$$\|\mathbf{x}\|_p^p = \sum_j |x_j|^p, \quad \mathbf{x} = (x_1, x_2, \dots, x_n),$$

show the inequality

$$\|\mathbf{x}+\mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p,$$

by looking at the optimization problem

$$\text{Minimize in } (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{N+N} : \quad (\|\mathbf{x}\|_p + \|\mathbf{y}\|_p)$$

subject to $\mathbf{x}+\mathbf{y}=\mathbf{z}$, for a fixed vector $\mathbf{z} \in \mathbb{R}^N$.

(a) Show that for every $\mathbf{z} \in \mathbb{R}^N$, there is always a unique optimal feasible pair $(\mathbf{x}, \mathbf{y})$.
(b) Use optimality to find such optimal pair.
(c) Conclude the desired inequality.

3. (a) Consider a variational problem of order zero of the form

$$\int_\Omega F(\mathbf{x}, u(\mathbf{x}))\, d\mathbf{x}.$$

Argue how to find the minimizer under no further constraint. Derive optimality conditions. Why can this same strategy not be implemented for a first-order problem?

(b) For a fixed function $f(\boldsymbol{x})$, consider the variational problem

$$\text{Minimize in } \chi(\boldsymbol{x}) : \quad \int_\Omega \chi(\boldsymbol{x}) f(\boldsymbol{x})\, d\boldsymbol{x}$$

for every characteristic function χ with

$$\int_\Omega \chi(\boldsymbol{x})\, d\boldsymbol{x} = |\Omega| s, \quad s \in (0, 1).$$

Describe the minimizers.

4. If the cost of flying in the air at height $h > 0$ over ground ($h = 0$) is given by e^{-ah}, write precisely the optimization problem providing the minimum cost of flying a distance L on ground level.
5. Fermat proposed in 1662 the principle of propagation of light: the path taken by light between two points is always the one minimizing the transit time. If the velocity of light in any medium is the quotient $c/\boldsymbol{n}$ where $\boldsymbol{n}(x, y)$ is the refractive index of the medium at point (x, y), and c is the velocity of light in vacuum, write the variational problem that furnishes the path followed by a ray of light going from point $(0, y_0)$ to (L, y_L).

6. Newton's solid of minimal resistance. The integral

$$\int \frac{y(x)y'(x)^3}{1+y'(x)^2}\,dx$$

provides the global resistance experienced by the solid of revolution obtained by the rotation of the graph $(x, y(x))$ around the X-axis between two points $(0, H)$, (L, h), $h < H$. Write the same previous functional for profiles described in the form $(x(y), y)$ between $x(H) = 0$, $x(h) = L$.
7. Minimal surfaces of revolution. One can restrict the minimal surface problem to the class of surfaces of revolution generated by graphs of functions $u(x)$ of a single variable that rotates around the X-axis between two values $x = 0$ and $x = L$. Write the functional to be minimized for such a problem.
8. Geodesics in a cylinder and in a sphere. Write the functional yielding the length of a curve σ whose image is contained either in the cylinder

$$x^2 + y^2 = 1$$

or in the sphere

$$x^2 + y^2 + z^2 = 1.$$

Discuss the several possibilities of the mutual position of the initial and final points in each case.
9. Consider the canonical parametrization $(x, y, u(x, y))$ of the graph of a smooth function

$$u(x, y), \quad (x, y) \in \Omega \subset \mathbb{R}^2.$$

Write the functional with integrand yielding the product of the sizes of the two generators of the tangent plane at each point of such a graph. Can you figure out at least one minimizer?
10. Let

$$\mathbf{u}(t) : [0, 1] \to \mathbb{R}^N$$

be a smooth map, and consider the functional

$$\int_0^1 F(\mathbf{u}(t), \mathbf{u}'(t))\,dt$$

for a certain integrand

$$F(\mathbf{u}, \mathbf{v}) : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}.$$

Write down the variational problem of finding the minimum of the above functional over all possible reparametrizations of the given path $\mathbf{u}$ preserving the values $\mathbf{u}(0)$ and $\mathbf{u}(1)$.

(a) How can the problem be simplified for smooth, bijective parametrizations?
(b) Deduce for which integrands $F(\mathbf{u}, \mathbf{v})$ the corresponding functional is independent of reparametrizations, and check that this is so for the one yielding the length of the curve $\{\mathbf{u}(t) : t \in [0, 1]\}$.

Part I
Basic Functional Analysis and Calculus of Variations

Chapter 2
A First Exposure to Functional Analysis

2.1 Overview

We focus on this chapter on the three main basic issues that are indispensable to tackle variational problems under a solid foundation, namely,

1. define the size of a function;
2. examine a compactness principle valid for bounded (with respect to that size) sequences of functions;
3. study whether limit functions enjoy similar properties to the members of a sequence.

From the very beginning, and as is usual in Mathematics, we will adopt an abstract viewpoint to gain as much generality as possible with the same effort. Fundamental spaces of functions and sequences will be explored as examples to illustrate concepts, fine points in results, counterexamples, etc. Spaces that are fundamental to Calculus of Variations and Partial Differential Equations (PDE, for short) will also be introduced as they will play a central role in our discussion. In particular, we will introduce the concept of weak derivative and weak solution of a PDE, and start studying the important Sobolev spaces.

From the strict perspective of Functional Analysis, we will introduce Banach and Hilbert spaces, the dual space of a given Banach space, weak topologies, and the crucial principle of weak compactness. Some other fundamental concepts of Functional Analysis will be deferred until later chapters.

2.2 Metric, Normed and Banach Spaces

Since sets of functions taking values on a vector space (over a certain field **K** which most of the time will be $\mathbb{R}$ if not explicitly claimed otherwise) inherit the structure

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4_2

of a vector space over the same field, we will assume that an abstract set $\mathbb{E}$ is such a vector space.

Definition 2.1 A non-negative function

$$\|\cdot\| : \mathbb{E} \to \mathbb{R}^+$$

is called a norm in $\mathbb{E}$ if:

N1 $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$;
N2 triangle inequality: for every pair $\mathbf{x}, \mathbf{y} \in \mathbb{E}$,

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|;$$

N3 for every scalar λ and vector $\mathbf{x} \in \mathbb{E}$,

$$\|\lambda \mathbf{x}\| = |\lambda| \, \|\mathbf{x}\|.$$

A norm is always a way to measure the size of vectors, and those three properties specify how it should behave with respect to linear combinations, the basic operation in a vector space.

If a set $\mathbb{E}$ is not a vector space, we can still define a distance function

$$d(\mathbf{x}, \mathbf{y}) : \mathbb{E} \times \mathbb{E} \to \mathbb{R}^+$$

to measure how far from each other elements of $\mathbb{E}$ are. Such distance function must comply with some basic properties to maintain a certain order in $\mathbb{E}$, namely:

1. $d(\mathbf{x}, \mathbf{y}) = 0$ only when $\mathbf{x} = \mathbf{y}$;
2. symmetry: for every pair $\mathbf{x}, \mathbf{y} \in \mathbb{E}$, $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$;
3. triangular inequality:

$$d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$$

for any three elements $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{E}$.

If $\|\cdot\|$ is a norm in $\mathbb{E}$, then the formula

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

yields a distance in $\mathbb{E}$, with the additional properties

1. $d(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z}) = d(\mathbf{x}, \mathbf{y})$;
2. $d(\lambda \mathbf{x}, \lambda \mathbf{y}) = |\lambda| d(\mathbf{x}, \mathbf{y})$.

Conversely, if

$$d(\cdot, \cdot) : \mathbb{E} \times \mathbb{E} \to \mathbb{R}^+$$

is a distance function enjoying the above two properties, then

$$\|\mathbf{x}\| = d(\mathbf{0}, \mathbf{x})$$

is a norm in $\mathbb{E}$. The topology in $\mathbb{E}$ is determined through this distance function. The collection of balls

$$\{\mathbf{x} \in \mathbb{E} : \|\mathbf{x}\| < r\}$$

for positive r make up a basis of neighborhoods at $\mathbf{0}$. Their translation to every vector $\mathbf{x} \in \mathbb{E}$ form a basis of neighborhoods at $\mathbf{x}$.

Definition 2.2

1. A complete metric space is a metric space $(\mathbb{E}, d)$ in which Cauchy sequences always converge to a point in the same space.
2. A Banach space $\mathbb{E}$ is a normed vector space which is complete with respect to its norm. Every Banach space is a complete metric space under the distance induced by the norm.

Example 2.1 The first example is, of course, that of a finite dimensional vector space $\mathbf{K}^N$ in which one can select various ways to measure the size of vectors:

$$\|\mathbf{x}\|_2^2 = \sum_i x_i^2, \quad \mathbf{x} = (x_1, x_2, \ldots, x_N),$$

$$\|\mathbf{x}\|_p^p = \sum_i |x_i|^p, \quad p \geq 1,$$

$$\|\mathbf{x}\|_\infty = \max_i |x_i|.$$

$\mathbf{K}^N$ is, definitely, complete if $\mathbf{K}$ is.

Example 2.2 One of the easiest, non-finite-dimensional normed spaces that can be shown to be complete is

$$\ell^\infty (= \ell^\infty(\mathbb{R})) = \{\mathbf{x} = (x_n)_{n \in \mathbb{N}} : \sup_n |x_n| = \|\mathbf{x}\|_\infty < \infty\}.$$

If a sequence

$$\{\mathbf{x}^{(j)}\} \subset \ell^\infty$$

is a Cauchy sequence, then

$$\|\mathbf{x}^{(j)} - \mathbf{x}^{(k)}\|_\infty \to 0 \quad \text{as } j, k \to \infty.$$

In particular, for every $n \in \mathbb{N}$,

$$|x_n^{(j)} - x_n^{(k)}| \le \|\mathbf{x}^{(j)} - \mathbf{x}^{(k)}\|_\infty.$$

Hence

$$\{x_n^{(j)}\} \subset \mathbb{R}$$

is a Cauchy sequence in $\mathbb{R}$, and consequently it converges to, say, x_n. It is then easy to check that

$$\mathbf{x} = (x_n)_{n \in \mathbb{N}}$$

belongs to ℓ^∞, and $\mathbf{x}^{(j)} \to \mathbf{x}$, that is to say

$$\|\mathbf{x}^{(j)} - \mathbf{x}\|_\infty \to 0 \text{ as } j \to \infty.$$

ℓ^∞ is not a finite-dimension vector space as it cannot admit a basis with a finite number of sequences.

Example 2.3 This is an extension of Example 2.2. Let $\mathbb{X}$ be a non-empty set, and take

$$\ell^\infty(\mathbb{X}) = \{f : \mathbb{X} \to \mathbb{R},\ f(\mathbb{X}) \text{ is bounded in } \mathbb{R}\},$$

and put

$$\|f\|_\infty = \sup\{|f(\mathbf{x})| : \mathbf{x} \in \mathbb{X}\}.$$

Just as above $(\ell^\infty(\mathbb{X}), \|\cdot\|_\infty)$ is a Banach space.

Example 2.4 Another fundamental example is that of the space of bounded, continuous functions $\mathcal{C}^0(\Omega)$ in an open set Ω of $\mathbb{R}^N$, under the sup-norm. In fact, if $\mathbb{X}$ is a general topological space, and $\mathcal{C}^0(\mathbb{X})$ denotes the space of bounded, continuous functions as a subspace of $\ell^\infty(\mathbb{X})$, because $\mathcal{C}^0(\mathbb{X})$ is a closed subspace of $\ell^\infty(\mathbb{X})$ (this is a simple exercise), it becomes a Banach space on its own under the sup-norm.

The most important examples in Applied Analysis are, by far, the L^p-spaces. As a prelude to their study, we look first at spaces ℓ^p for $p \ge 1$, in a similar way as with ℓ^∞.

Example 2.5 We first define

$$\ell^p = \{\mathbf{x} = (x_n)_{n\in\mathbb{N}} : \|\mathbf{x}\|_p^p \equiv \sum_n |x_n|^p < \infty\}, \quad p > 0.$$

It turns out that ℓ^p is a Banach space precisely because

$$\|\mathbf{x}\|_p = \left(\sum_n |x_n|^p\right)^{1/p}, \quad p \geq 1,$$

is a norm in the vector space of sequences of real numbers without any further condition. The triangular inequality was checked in an exercise in the last chapter. The other two conditions are trivial to verify. Hence ℓ^p, for exponent $p \geq 1$, becomes a normed-vector space. The reason why it is complete is exactly as in the ℓ^∞ case. Simply note that

$$|x_n^{(j)} - x_n^{(k)}| \leq \|\mathbf{x}^{(j)} - \mathbf{x}^{(k)}\|_p, \quad \mathbf{x}^{(j)} = (x_n^{(j)}), \mathbf{x}^{(k)} = (x_n^{(k)}),$$

for every n. Hence if $\{\mathbf{x}^{(j)}\}$ is a Cauchy sequence in ℓ^p, each component is a Cauchy sequence of real numbers. The case $p = 2$ is very special as we will later see. For exponent $p \in (0, 1)$ the sets ℓ^p are not, in general, vector spaces. See Exercise 26 below.

2.3 Completion of Normed Spaces

Every normed space can be densely embedded in a Banach space. Every metric space can be densely embedded in a complete metric space. The process is similar to the passage from the rationals to the reals. Recall Example 2.3 above.

Theorem 2.1

1. *Let* $(\mathbb{E}, d)$ *be a metric space. A complete metric space* $(\hat{\mathbb{E}}, \hat{d})$, *and an isometry*

$$\Phi : (\mathbb{E}, d) \mapsto (\hat{\mathbb{E}}, \hat{d})$$

   can be found such that $\Phi(\mathbb{E})$ *is dense in* $\hat{\mathbb{E}}$. $(\hat{\mathbb{E}}, \hat{d})$ *is unique modulus isometries.*
2. *Every normed space* $(\mathbb{E}, \|\cdot\|)$ *can be densely embedded in a Banach space* $(\hat{\mathbb{E}}, \|\cdot\|)$ *through a linear isometry*

$$\Phi : (\mathbb{E}, \|\cdot\|) \mapsto (\hat{\mathbb{E}}, \|\cdot\|).$$

Proof Choose a point $\mathbf{p} \in \mathbb{E}$, and define

$$\Phi : (\mathbb{E}, d) \mapsto \ell^\infty(\mathbb{E}), \quad \Phi(\mathbf{q})(\mathbf{x}) = d(\mathbf{q}, \mathbf{x}) - d(\mathbf{p}, \mathbf{x}).$$

Because of the triangular inequality, it is true that

$$|d(\mathbf{q}, \mathbf{x}) - d(\mathbf{p}, \mathbf{x})| \leq d(\mathbf{q}, \mathbf{p})$$

for every $\mathbf{x} \in \mathbb{E}$, and

$$\|\Phi(\mathbf{q})\|_\infty \leq d(\mathbf{q}, \mathbf{p}).$$

In a similar way, for every pair $\mathbf{q}_i$, $i = 1, 2$,

$$\|\Phi(\mathbf{q}_1) - \Phi(\mathbf{q}_2)\|_\infty \leq d(\mathbf{q}_1, \mathbf{q}_2).$$

However

$$|\Phi(\mathbf{q}_1)(\mathbf{q}_2) - \Phi(\mathbf{q}_2)(\mathbf{q}_2)| = d(\mathbf{q}_1, \mathbf{q}_2)$$

and, hence,

$$\|\Phi(\mathbf{q}_1) - \Phi(\mathbf{q}_2)\|_\infty = d(\mathbf{q}_1, \mathbf{q}_2).$$

Take

$$(\hat{\mathbb{E}}, \hat{d}) = (\overline{\Phi(\mathbb{E})}, \|\cdot\|_\infty)$$

where closure is meant in the Banach space $\ell^\infty(\mathbb{E})$. The uniqueness is straightforward. In fact, suppose that

$$\Phi_i : (\mathbb{E}, d) \mapsto \Phi_i(\mathbb{E}) \subset (\mathbb{E}_i, d_i), \quad i = 1, 2,$$

with $\Phi_i(\mathbb{E})$ dense in $\mathbb{E}_i$. The mapping

$$\Phi = \Phi_2 \circ \Phi_1^{-1} : \Phi_1(\mathbb{E}) \mapsto \Phi_2(\mathbb{E})$$

can be extended, by completeness and because it is an isometry, to $\mathbb{E}_1$ and put

$$\Phi : (\mathbb{E}_1, d_1) \mapsto (\mathbb{E}_2, d_2)$$

so that

$$\Phi|_{\Phi_1(\mathbb{E})} = \Phi_2 \circ \Phi_1^{-1};$$

and viceversa, there is

$$\Theta : (\mathbb{E}_2, d_2) \mapsto (\mathbb{E}_1, d_1), \quad \Theta|_{\Phi_2(\mathbb{E})} = \Phi_1 \circ \Phi_2^{-1}.$$

Therefore

$$\Phi \circ \Theta = \mathbf{1}|_{\mathbb{E}_2}, \quad \Theta \circ \Phi = \mathbf{1}|_{\mathbb{E}_1},$$

in respective dense subsets. The claim is proved, then, through completeness.

Since $(\mathbb{E}, \|\cdot\|)$ can be understood as a metric space under the distance

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|,$$

the second part is a consequence of the first. Let

$$\Phi : (\mathbb{E}, \|\cdot\|) \mapsto (\hat{\mathbb{E}}, \hat{d})$$

be such an isometry with $\Phi(\mathbb{E})$ densely embedded in $\hat{\mathbb{E}}$. If $\mathbf{x}$, $\mathbf{y}$ are elements of $\hat{\mathbb{E}}$, one can take Cauchy sequences $\{\mathbf{x}_j\}$, $\{\mathbf{y}_j\}$ in $\mathbb{E}$ such that

$$\Phi(\mathbf{x}_j) \to \mathbf{x}, \quad \Phi(\mathbf{y}_j) \to \mathbf{y}.$$

In that case $\{\mathbf{x}_j + \mathbf{y}_j\}$ is also a Cauchy sequence, and by the completeness of $\hat{E}$, $\Phi(\mathbf{x}_j + \mathbf{y}_j) \to \mathbf{x} + \mathbf{y}$. The same is correct for the product by scalars. It is straightforward to check the independence of these manipulations from the chosen Cauchy sequences. Finally,

$$\|\mathbf{x}\| = \hat{d}(\mathbf{x}, \mathbf{0}) = \lim_{j\to\infty} \hat{d}(\Phi(\mathbf{x}_j), \mathbf{0}) = \lim_{j\to\infty} \|\mathbf{x}_j\|$$

and

$$\hat{d}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

□

This abstract result informs us that every time we have a norm on a vector space, we can automatically consider its completion with respect to that norm. The resulting space becomes a Banach space in which the starting one is dense. This is often the way to tailor new spaces for specific purposes.

2.4 L^p-Spaces

Let $\Omega \subset \mathbb{R}^N$ be an open subset. $d\mathbf{x}$ will indicate the Lebesgue measure in $\mathbb{R}^N$. We assume that readers have a sufficient basic background on Measure Theory to know what it means for a function

$$f(\mathbf{x}) : \Omega \to \mathbb{R}$$

to be measurable and integrable, and to have a finite essential supremum. All functions considered are measurable without further notice. In fact, it is usually written

$$L^1(\Omega) = \{f(\mathbf{x}) : \Omega \to \mathbb{R} : f \text{ is integrable}\},$$

and define

$$\|f\|_1 = \int_\Omega |f(\mathbf{x})| \, d\mathbf{x}.$$

Definition 2.3 Let $p \geq 1$. We put

$$L^p(\Omega) = \{f(\mathbf{x}) : \Omega \to \mathbb{R} : |f(\mathbf{x})|^p \text{ is integrable}\},$$

and

$$L^\infty(\Omega) = \{f(\mathbf{x}) : \Omega \to \mathbb{R} : \text{esssup}_{\mathbf{x}\in\Omega} |f(\mathbf{x})| < +\infty\}.$$

The main objective of this section is to show that $L^p(\Omega)$ is a Banach space for every $p \in [1, +\infty]$ with norm

$$\|f\|_p^p = \int_\Omega |f(\mathbf{x})|^p \, d\mathbf{x}, \quad \|f\|_\infty = \text{esssup}_{\mathbf{x}\in\Omega} |f(\mathbf{x})|. \tag{2.1}$$

This goal proceeds in two main steps:

1. show that $L^p(\Omega)$ is a vector space, and $\|\cdot\|_p$ is a norm; and
2. $L^p(\Omega)$ is complete under this p-th norm.

It is trivial to check that

$$\|\lambda f\|_p = |\lambda| \|f\|_p$$

for every scalar $\lambda \in \mathbb{R}$, and every $f \in L^p(\Omega)$. More involved is to ensure that

$$f + g \in L^p(\Omega)$$

as soon as both f and g are functions in $L^p(\Omega)$. But in fact, this is a direct consequence of the fact that $\|\cdot\|_p$ is a norm; more in particular, of the triangle inequality N2 in Definition 2.23. Indeed, given a large vector space $\mathbb{E}$, if $\|\cdot\|$ is a norm over $\mathbb{E}$ that can take on the value $+\infty$ some times, then

$$\mathbb{E}_{\|\cdot\|} = \{\mathbf{x} \in \mathbb{E} : \|\mathbf{x}\| < +\infty\}$$

is a vector subspace of $\mathbb{E}$. All we have to care about, then, is that the p-th norm given above is indeed a norm in the space of measurable functions over Ω. Both N1 and N3 are trivial. We focus on N2. This is in turn a direct consequence of an important inequality for numbers.

Lemma 2.1 (Young's Inequality) *For real numbers a and b, and exponent $p > 1$, we have*

$$|ab| \le \frac{1}{p}|a|^p + \frac{1}{q}|b|^q, \qquad \frac{1}{p} + \frac{1}{q} = 1.$$

Proof It is elementary to check that the function $\log x$ is concave for x, positive. This exactly means that

$$t_1 \log a_1 + t_1 \log a_2 \le \log(t_1 a_1 + t_2 a_2), \qquad t_i \ge 0, t_1 + t_2 = 1, a_i > 0.$$

For the choice

$$t_1 = \frac{1}{p}, t_2 = \frac{1}{q}, \qquad a_1 = |a|^p, a_2 = |b|^q,$$

we find, through the basic properties of logarithms, that

$$\log|ab| = \frac{1}{p}\log|a|^p + \frac{1}{q}\log|b|^q \le \log\left(\frac{1}{p}|a|^p + \frac{1}{q}|b|^q\right).$$

The increasing monotonicity of the logarithmic function implies the claimed inequality. □

Another necessary, but fundamental inequality on its own right, follows. It is a consequence of the previous one.

Lemma 2.2 (Hölder's Inequality) *If $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$, i.e.*

$$\|f\|_p < \infty, \qquad \|g\|_q < \infty,$$

with $1 = 1/p + 1/q$, then the product fg is integrable ($fg \in L^1(\Omega)$), and

$$\|fg\|_1 \le \|f\|_p \|g\|_q.$$

Proof Apply Young's inequality to the choice

$$a = \frac{|f(\mathbf{x})|}{\|f\|_p}, \quad b = \frac{|g(\mathbf{x})|}{\|g\|_q},$$

for an arbitrary point $\mathbf{x} \in \Omega$, to obtain

$$\frac{|f(\mathbf{x})||g(\mathbf{x})|}{\|f\|_p \|g\|_q} \le \frac{1}{p} \frac{|f(\mathbf{x})|^p}{\|f\|_p^p} + \frac{1}{q} \frac{|g(\mathbf{x})|^q}{\|g\|_q^q}.$$

Rearranging terms and integrating over Ω, we arrive at our conclusion. □

We are now ready to show the following.

Proposition 2.1 *For $p \in [1, +\infty]$, the p-th norm in (2.1) is a norm in the space of measurable functions in Ω, and consequently $L^p(\Omega)$ is a vector space.*

Proof The limit cases $p = 1$ and $p = \infty$ are straightforward (left as an exercise). Take $1 < p < \infty$, and let f and g be two measurable functions defined in Ω. Then for each individual point $\mathbf{x} \in \Omega$, we have

$$|f(\mathbf{x}) + g(\mathbf{x})|^p \le |f(\mathbf{x}) + g(\mathbf{x})|^{p-1}|f(\mathbf{x})| + |f(\mathbf{x}) + g(\mathbf{x})|^{p-1}|g(\mathbf{x})|.$$

We apply Hölder's inequality to both terms to find

$$\|f + g\|_p^p \le \||f + g|^{p-1}\|_q \|f\|_p + \||f + g|^{p-1}\|_q \|g\|_p. \tag{2.2}$$

If we note that $q = p/(p-1)$, then it is immediate to check that

$$\||f + g|^{p-1}\|_q = \|f + g\|_p^{p-1},$$

and taking this identity to (2.2), we conclude the proof. □

The second step is equally important.

Proposition 2.2 *$L^p(\Omega)$ is a Banach space for all $p \in [1, +\infty]$, and every open subset $\Omega \subset \mathbb{R}^N$.*

Proof The case $p = \infty$ was already examined in the previous section. Take a finite $p \ge 1$, and suppose that $\{f_j\}$ is a Cauchy sequence in the vector space $L^p(\Omega)$, i.e.

$$\|f_j - f_k\|_p \to 0 \text{ as } j, k \to \infty.$$

It is then easy to argue that for a.e. $\mathbf{x} \in \Omega$,

$$|f_j(\mathbf{x}) - f_k(\mathbf{x})| \to 0 \text{ as } j, k \to \infty. \tag{2.3}$$

Indeed, if for each fixed $\epsilon > 0$, we let

$$\Omega_{\epsilon,j,k} \equiv \{\mathbf{x} \in \Omega : |f_j(\mathbf{x}) - f_k(\mathbf{x})| \geq \epsilon\},$$

then

$$\lim_{j,k\to\infty} |\Omega_{\epsilon,j,k}| = 0.$$

If, seeking a contradiction, this were not true, then for some $\epsilon > 0$, we could find $\delta > 0$, such that

$$\lim_{j,k\to\infty} |\Omega_{\epsilon,j,k}| \geq \delta.$$

But then

$$\begin{aligned} 0 < \delta\epsilon^p \leq & \lim_{j,k\to\infty} \int_{\Omega_{\epsilon,j,k}} |f_j(\mathbf{x}) - f_k(\mathbf{x})|^p \, d\mathbf{x} \\ \leq & \int_\Omega |f_j(\mathbf{x}) - f_k(\mathbf{x})|^p \, d\mathbf{x} \\ = & \|f_j - f_k\|_p^p, \end{aligned}$$

which is a contradiction. Equation (2.3) is therefore correct, and there is a limit function $f(\mathbf{x})$. For k, fixed but arbitrary, set

$$g_j(\mathbf{x}) = |f_j(\mathbf{x}) - f_k(\mathbf{x})|^p.$$

Each g_j is integrable, non-negative, and

$$\int_\Omega g_j(\mathbf{x}) \, d\mathbf{x} = \int_\Omega |f_j(\mathbf{x}) - f_k(\mathbf{x})|^p \, d\mathbf{x}.$$

Hence $\{g_j\}$ is bounded in $L^1(\Omega)$. By the classical Fatou's lemma, we can conclude that

$$\int_\Omega |f(\mathbf{x}) - f_k(\mathbf{x})|^p \, d\mathbf{x} \leq \liminf_{j\to\infty} \|f_j - f_k\|_p^p$$

for all such k. Taking limits in k in both sides of this inequality, we deduce that

$$\limsup_{k\to\infty} \int_\Omega |f(\mathbf{x}) - f_k(\mathbf{x})|^p \, d\mathbf{x} \leq \lim_{j,k\to\infty} \|f_j - f_k\|_p^p = 0.$$

Then

$$\|f - f_k\|_p \to 0,$$

and f is the limit function. □

There are many more important properties of these fundamental spaces, some of which will be stated and proved later.

2.5 Weak Derivatives

Our intention of not loosing sight of variational principles pushes us to move beyond L^p-spaces, as we need to cope with derivatives as an essential ingredient. To be more specific, suppose we consider the one-dimensional functional

$$I(u) = \int_{x_0}^{x_1} F(u(x), u'(x))\, dx \tag{2.4}$$

for competing functions

$$u(x) : [x_0, x_1] \subset \mathbb{R} \to \mathbb{R},$$

where the integrand

$$F(u, z) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

is supposed smooth and regular. At first sight, as soon as we write the derivative $u'(x)$, we seem to have no alternative than to restrict attention to $\mathcal{C}^1$-functions. Otherwise, we do not know what the derivative $u'(x)$ might mean in those points where u is not differentiable. The truth is that because the value of the integrals in (2.4) is determined uniquely when the integrand

$$F(u(x), u'(x))$$

is defined except, possibly, in a finite number of points, we realize that we can permit piecewise, $\mathcal{C}^1$ functions to enter into the optimization process. Indeed, we can allow many more functions.

It all starts with the well-known integration-by-parts formula that is taught in elementary Calculus courses. That formula, in turn, is a direct consequence of the product rule for differentiation, and the Fundamental Theorem of Calculus. It reads

$$\int_a^b f'(x)g(x)\, dx = f(x)g(x)|_{x=a}^{x=b} - \int_a^b f(x)g'(x)\, dx. \tag{2.5}$$

In such basic courses, one is told that this formula is valid whenever both functions f and g are continuously differentiable in the interval $[a, b]$. This is indeed so. Suppose, however, that f is only continuous in $[a, b]$ but fails to be differentiable at some points, and that we could, somehow, find another function F, not even continuous, just integrable, in such a way that

$$\int_a^b F(x)g(x)\,dx = -\int_a^b f(x)g'(x)\,dx \tag{2.6}$$

holds for every continuously differentiable function g with $g(a) = g(b) = 0$.

Example 2.6 Take

$$a = -1, \quad b = 1, \quad f(x) = |x|$$

which is not differentiable at the origin, and

$$F(x) = \begin{cases} -1, & -1 \le x < 0, \\ 1, & 0 < x \le 1. \end{cases}$$

We will take $g(x)$ as an arbitrary continuously differentiable function with

$$g(-1) = g(1) = 0.$$

Because integrals enjoy the additivity property with respect to intervals of integration, we can certainly write

$$\int_{-1}^1 F(x)g(x)\,dx = \int_{-1}^0 F(x)g(x)\,dx + \int_0^1 F(x)g(x)\,dx,$$

and then

$$\int_{-1}^0 F(x)g(x)\,dx = -\int_{-1}^0 g(x)\,dx,$$
$$\int_0^1 F(x)g(x)\,dx = \int_0^1 g(x)\,dx.$$

On the other hand,

$$\int_{-1}^1 f(x)g'(x)\,dx = \int_{-1}^0 -xg'(x)\,dx + \int_0^1 xg'(x)\,dx,$$

and a "true" integration by parts in these two integrals separately clearly yields

$$\int_{-1}^{1} f(x)g'(x)\,dx = \int_{-1}^{0} g(x)\,dx - \int_{0}^{1} g(x)\,dx = -\int_{-1}^{1} F(x)g(x)\,dx.$$

We therefore see that formula (2.6) is formally correct if we put $f'(x) = F(x)$. Notice that the difficulty at $x = 0$ does not have any relevance in these computations because what happens at a single point (a set of measure zero) is irrelevant for integrals.

After this example, we suspect that formula (2.6) hides a more general concept of derivative.

Definition 2.4 Let

$$f(x) : (a, b) \to \mathbb{R}$$

be a measurable, integrable function. Another such function

$$F(x) : (a, b) \to \mathbb{R}$$

is said to be the weak derivative of f, and we write $f' = F$, if (2.6) holds for every continuously differentiable function

$$g(x) : [a, b] \to \mathbb{R}, \text{ with } g(a) = g(b) = 0.$$

In this way, we would say that $F(x) = x/|x|$ is the weak derivative of $f(x) = |x|$ even if F is not defined at $x = 0$, and f is not differentiable at the same point. We can conclude that formula (2.6) is always correct as long as the derivative of f is understood in a weak sense. Of course, if f is continuously differentiable, then its standard derivative is also its weak derivative.

Example 2.7 Even though the function $f(x) = \sqrt{|x|}$ has a derivative

$$f'(x) = \frac{x}{2|x|\sqrt{|x|}}, \qquad x \neq 0,$$

that is not defined at $x = 0$, and, in fact, it blows up at this point, this function $f'(x)$ is also the weak derivative of f in the full interval $[-1, 1]$. Indeed, it is easy to check that the formula of integration by parts

$$\int_{-1}^{1} f'(x)g(x)\,dx = -\int_{-1}^{1} f(x)g'(x)\,dx$$

is correct for every $\mathcal{C}^1$-function g vanishing at $x = \pm 1$.

2.6 One-Dimensional Sobolev Spaces

We are ready to introduce the fundamental Sobolev spaces of weakly differentiable functions that are so indispensable for PDEs and Calculus of Variations. We will restrict attention now to the one-dimensional situation, and deferred the treatment of the more involved higher-dimensional case.

Definition 2.5 Let $J = (a, b) \subset \mathbb{R}$ be a finite interval, and take $1 \le p \le \infty$. The Sobolev space $W^{1,p}(J)$ is the subspace of functions f of $L^p(J)$ with a weak derivative f' which is also a function in $L^p(J)$, namely

$$W^{1,p}(J) = \{f \in L^p(J) : f' \in L^p(J)\}.$$

The first thing to do is to determine the norm in $W^{1,p}(J)$, and ensure that $W^{1,p}(J)$ is indeed a Banach space.

Proposition 2.3 *For J and p as above, $W^{1,p}(J)$ is a Banach space under the norm*

$$\|f\|_{1,p} = \|f\|_p + \|f'\|_p.$$

Equivalently, we can also take

$$\|f\|_{1,p}^p = \int_J [|f(x)|^p + |f'(x)|^p]\, dx.$$

Proof The fact that $\|\cdot\|_{1,p}$ is a norm in $W^{1,p}(J)$ is a direct consequence of the fact that $\|\cdot\|_p$ is in $L^p(\Omega)$. To show that $W^{1,p}(J)$ is a Banach space is now easy, given that convergence in the $\|\cdot\|_{1,p}$ means convergence in $L^p(J)$ of functions and derivatives. In fact,

$$\|f_j - f_k\|_{1,p} \to 0 \quad \text{as } j, k \to \infty,$$

is equivalent to

$$\|f_j - f_k\|_p, \|f_j' - f_k'\|_p \to 0 \quad \text{as } j, k \to \infty.$$

There are functions $f, g \in L^p(J)$ with

$$f_j \to f, \quad f_j' \to g$$

in $L^p(J)$. It remains to check that $g = f'$. To this end, take an arbitrary, $\mathcal{C}^1$-function

$$\phi(x) : J \to \mathbb{R}, \quad \phi(a) = \phi(b) = 0.$$

From

$$\int_J f_j(x)\phi'(x)\,dx = -\int_J f_j'(x)\phi(x)\,dx,$$

by taking limits in j, we find (justification: left as an exercise)

$$\int_J f(x)\phi'(x)\,dx = -\int_J g(x)\phi(x)\,dx.$$

The arbitrariness of ϕ in this identity, exactly means that $f' = g$, hence

$$f \in W^{1,p}(J), \quad f_j \to f$$

in the Sobolev space. □

In dimension one, there is essentially a unique source of functions in $W^{1,p}(J)$. It is a consequence of the Fundamental Theorem of Calculus.

Lemma 2.3 *Let $g \in L^p(J)$, and put*

$$f(x) = \int_a^x g(s)\,ds.$$

Then $f \in W^{1,p}(J)$, and $f' = g$.

Proof We first check that f, so defined, belongs to $L^p(J)$. Let $p \geq 1$ be finite. Note that for each $x \in J$, by Hölder's inequality applied to the two factors $g \in L^p(J)$ and the characteristic function of the interval of integration,

$$|f(x)| \leq \int_a^x |g(s)|\,ds \leq \|g\|_p |x-a|^{(p-1)/p} \leq \|g\|_p |J|^{(p-1)/p}.$$

The arbitrariness of $x \in J$ shows that $f \in L^\infty(J)$, and, in particular, $f \in L^p(J)$ for every p.

Let us check that the formula of integration by parts is correct. So take a $\mathcal{C}^1$-function $\phi(x)$ with $\phi(a) = \phi(b) = 0$. Then

$$\int_a^b f(x)\phi'(x)\,dx = \int_a^b \int_a^x g(s)\phi'(x)\,ds\,dx.$$

Fubini's theorem permits us to interchange the order of integration in this double integral to find that

$$\int_a^b f(x)\phi'(x)\,dx = \int_a^b \int_s^b g(s)\phi'(x)\,dx\,ds.$$

If we change the name of the dummy variables of integration to avoid any confusion, we can also write

$$\int_a^b f(x)\phi'(x)\,dx = \int_a^b \int_x^b g(x)\phi'(s)\,ds\,dx,$$

and the Fundamental Theorem of Calculus applied to the smooth function ϕ (recall that $\phi(b) = 0$) yields

$$\int_a^b f(x)\phi'(x)\,dx = -\int_a^b g(x)\phi(x)\,dx.$$

The arbitrariness of ϕ implies that indeed

$$f' = g \in L^p(J), \quad f \in W^{1,p}(J).$$

□

A funny consequence of this lemma amounts to having the validity of the Fundamental Theorem of Calculus for functions in $W^{1,p}(J)$.

Corollary 2.1 *Let $f \in W^{1,p}(J)$. Then for every point $c \in J$,*

$$f(x) - f(c) = \int_c^x f'(s)\,ds.$$

2.6.1 Basic Properties

It is also important to study some basic properties of these spaces. In particular, we would like to learn to appreciate how much more regular functions in $W^{1,p}(J)$ are compared to functions in $L^p(J) \setminus W^{1,p}(J)$. One main point from the perspective of the variational problems we would like to study is to check how end-point constraints like

$$u(x_0) = u_0, \quad u(x_1) = u_1$$

for given values $u_0, u_1 \in \mathbb{R}$ can be enforced. Note that such a condition is meaningless for functions in $L^p(J)$, $J = (x_0, x_1)$, because functions in L^p-spaces are defined except in negligible subsets, and individual points x_0 and x_1 do have measure zero. As usual J is a bounded interval either (a, b) or (x_0, x_1).

Proposition 2.4 *Let $p \in [1, \infty]$. Every function in $W^{1,p}(J)$ is absolutely continuous. If $p > 1$, every bounded set in $W^{1,p}(J)$ is equicontinuous.*

Proof Our starting point is the Fundamental Theorem of Calculus Corollary 2.1

$$f(x) - f(y) = \int_y^x f'(s)\,ds, \quad |f(x) - f(y)| \le \int_y^x |f'(s)|\,ds. \tag{2.7}$$

If we use Hölder's inequality in the integral in the right-hand side for the factors $|f'|$ and the characteristic function of the interval of integration as before, we deduce that

$$|f(x) - f(y)| \le \|f'\|_p |x - y|^{(p-1)/p}. \tag{2.8}$$

This inequality clearly implies that every uniformly bounded set (in particular a single function) in $W^{1,p}(J)$ is equicontinuous, if $p > 1$. For the case $p = 1$, the inequality breaks down because the size of $|x - y|$ is lost on the right-hand side. However, (2.7) still implies that a single function in $W^{1,1}(J)$ is absolutely continuous (why is the argument not valid for an infinite number of functions in a uniformly bounded set in $W^{1,1}(J)$?). Inequality (2.8) is referred to as the fact that a bounded sequence of functions in $W^{1,p}(J)$ with $1 < p < \infty$ is uniformly Hölder continuous with exponent

$$\alpha = 1 - 1/p \in (0, 1).$$

For $p = \infty$, the exponent is $\alpha = 1$, and we say that the set is uniformly Lipschitz.

□

As a consequence of this proposition, we can define the following important subspace of $W^{1,p}(J)$, because functions in $W^{1,p}(J)$ are continuous, and so its values at individual points are well-defined.

Definition 2.6 The space $W_0^{1,p}(J)$ is the subspace of functions of $W^{1,p}(J)$ whose end-point values vanish

$$W_0^{1,p}(J) = \{u \in W^{1,p}(J) : u(a) = u(b) = 0\}, \quad J = (a, b).$$

Proposition 2.4, together with the classic Arzelá-Ascoli theorem, implies that every bounded sequence in $W^{1,p}(J)$, with $p > 1$, admits a subsequence converging, in the $L^\infty(J)$-norm, to some function in this space. This deserves further closer attention.

2.6.2 Weak Convergence

Suppose that

$$f_j \to f \text{ in } L^\infty(J), \quad f_j \in W^{1,p}(J),\ f \in L^\infty(J). \tag{2.9}$$

In this situation, there are two issues that we would like to clarify:

1. Is it true that in fact $f \in W^{1,p}(J)$, i.e. there is some $f' \in L^p(J)$ so that

$$\int_J f(x)\phi'(x)\,dx = -\int_J f'(x)\phi(x)\,dx$$

   for every $\phi \in \mathcal{C}_0^1(J)$?[1]
2. If so, how is the convergence of f_j' to f'?

Under (2.9), we can write

$$\int_J f_j'(x)\phi(x)\,dx = -\int_J f_j(x)\phi'(x)\,dx \to -\int_J f(x)\phi'(x)\,dx. \tag{2.10}$$

Since this is correct for arbitrary functions ϕ in $\mathcal{C}_0^1(J)$, one can look at ϕ as a functional variable, and regard all those integrals as "linear operations" on ϕ. In particular, Hölder's inequality implies that

$$\left|\int_J f_j'(x)\phi(x)\,dx\right| \le \|f_j'\|_p\|\phi\|_q, \quad \frac{1}{p}+\frac{1}{q}=1.$$

The boundedness of $\{f_j\}$ in $W^{1,p}(J)$ indicates that the sequence of linear operations

$$\langle T_j, \phi\rangle = \int_J f_j'(x)\phi(x)\,dx$$

is such that

$$|\langle T_j, \phi\rangle| \le C\|\phi\|_q, \tag{2.11}$$

with C independent of j. This situation somehow forces us to bother about sequences of linear operations uniformly bounded.

Suppose one could argue somehow that, under (2.11), there is some function $g \in L^p(J)$ with

$$\langle T_j, \phi\rangle = \int_J f_j'(x)\phi(x)\,dx \to \int_J g(x)\phi(x)\,dx.$$

(2.10) would imply then that

$$\int_J g(x)\phi(x) = -\int_J f(x)\phi'(x)\,dx$$

[1] The notation $_0$ means that functions in this space vanish on end-points of J.

for all $\phi \in C_0^1(J)$, and this would precisely imply that indeed

$$f \in W^{1,p}(J), \quad f' = g.$$

Once we know this, the Fundamental Theorem of Calculus Corollary 2.1 leads to the fact that

$$\int_x^y f_j'(s)\,ds = f_j(x) - f_j(y) \to f(x) - f(y) = \int_x^y f'(s)\,ds,$$

for arbitrary points x and y in J. This suggests that the convergence at the level of derivatives f_j' to f' is such that

$$\int_x^y f_j'(s)\,ds \to \int_x^y f'(s)\,ds$$

for arbitrary points $x, y \in J$. This is indeed a notion of convergence as good as any other, but definitely different from the usual point-wise convergence as the following example shows.

Example 2.8 We recover Example 1.1 where

$$u_j(x) = \sin(2^j \pi x), \quad x \in [0, 1].$$

We already argued that there is no way to find a function $u(x)$ so that

$$u_j(x) \to u(x)$$

for a.e. $x \in [0, 1]$, not even for arbitrary subsequences. Yet, it is not difficult to check that

$$\int_a^b u_j(x)\,dx \to 0 = \int_a^b 0\,dx$$

for every arbitrary pair of points $a < b$ in $[0, 1]$ (left as an exercise). What this fact means is that the trivial, identically vanishing function is, somehow, determined by the sequence $\{u_j\}$ by passage to the limit in a special way. This is called weak convergence or convergence in the mean. Note that what makes this possible is the persistent cancellation of the positive and negative parts of u_j when computing the integral. In other words

$$\left|\int_a^b u_j(x)\,dx\right| \to 0 \text{ as } j \to \infty;$$

while it is not true that

$$\int_a^b |u_j(x)|\,dx \to 0.$$

Example 2.9 It is important to start realizing why the exponent $p = 1$ in Lebesgue and Sobolev spaces is so special. The main issue, from the scope of our interest in variational problems, that we would like to emphasize refers to weak compactness. We can understand the difficulty with a very simple example. Take

$$u_j(x) = j\chi_{[0,1/j]}(x) = \begin{cases} j, & 0 \le x \le 1/j \\ 0, & else \end{cases},$$

and put

$$U_j(x) = \int_0^x u_j(y)\,dy.$$

A simple computation and picture convince us that the sequence $\{U_j\}$ is uniformly bounded in $W^{1,1}(0, 1)$, and yet there cannot be a limit in whatever reasonable sense we may define it, because it would be sensible to pretend that the function $U \equiv 1$ be such limit but then we would have

$$0 = \int_0^1 U'(t)\,dt, \qquad \lim_{j\to\infty} \int_0^1 U_j'(x)\,dx = 1,$$

which would be pretty weird for a limit function U. The phenomenon happening with the sequence of derivatives $\{u_j\}$ is informally referred to as a concentration phenomenon at $x = 0$: derivatives concentrate its full finite "mass" in a smaller and smaller region around $x = 0$, as j becomes larger and larger; this concentration of mass forces a breaking of continuity at that point.

We see that we are forced to look more closely into linear operations on Banach spaces. This carries us directly into the dual space.

2.7 The Dual Space

This section is just a first incursion on the relevance of the dual space.

Definition 2.7 Let $\mathbb{E}$ be a Banach space. The dual space $\mathbb{E}'$ is defined to be the collection of all linear and continuous functionals defined on $\mathbb{E}$

$$\mathbb{E}' = \{T : \mathbb{E} \to \mathbb{R} : T, \text{ linear and continuous}\}.$$

A clear criterium to decide whether a linear operation is continuous follows.

Proposition 2.5 *A linear functional*

$$T : \mathbb{E} \to \mathbb{R}$$

belongs to $\mathbb{E}'$ *if and only if there is* $C > 0$ *with*

$$|T(\mathbf{x})| \leq C\|\mathbf{x}\| \tag{2.12}$$

for all $\mathbf{x} \in \mathbb{E}$.

Proof On the one hand, if $\mathbf{x}_j \to \mathbf{x}$ in $\mathbb{E}$, then

$$|T\mathbf{x}_j - T\mathbf{x}| = |T(\mathbf{x}_j - \mathbf{x})| \leq C\|\mathbf{x}_j - \mathbf{x}\| \to 0$$

as $j \to \infty$, and T is continuous. On the other, put

$$C = \sup_{\|\mathbf{x}\|=1} |T(\mathbf{x})|, \tag{2.13}$$

and suppose that $C = +\infty$, so that there is a certain sequence $\mathbf{x}_j$ with

$$\|\mathbf{x}_j\| = 1, \quad T(\mathbf{x}_j) \to \infty.$$

In this case

$$\frac{1}{T(\mathbf{x}_j)}\mathbf{x}_j \to \mathbf{0}$$

and yet, due to linearity,

$$T\left(\frac{1}{T(\mathbf{x}_j)}\mathbf{x}_j\right) = 1,$$

which is a contradiction with the continuity of T. The constant C in (2.13) is then finite, and, again through linearity, the estimate (2.12) is correct. □

This proposition is the basis of the natural possibility of regarding $\mathbb{E}'$ as a Banach space on its own right by putting

$$\|T\| = \sup_{\|\mathbf{x}\|=1} |T(\mathbf{x})|$$

or, in an equivalent way,

$$\|T\| = \inf\{M : |T(\mathbf{x})| \leq M\|\mathbf{x}\| \text{ for all } \mathbf{x} \in \mathbb{E}\}.$$

Proposition 2.6 *If $\mathbb{E}$ is a normed space (not necessarily complete), then $\mathbb{E}'$ is a Banach space endowed with this norm.*

The proof of this fact is elementary, and is left as an exercise.

Quite often, we write

$$T(\mathbf{x}) = \langle T, \mathbf{x}\rangle,$$

and refer to it as the duality pair or product, to stress the symmetric roles played by T and $\mathbf{x}$ in such an identity. As a matter of fact, if $\mathbf{x}' \in \mathbb{E}'$ is an arbitrary element, the duality pair and the corresponding estimate

$$\langle \mathbf{x}', \mathbf{x}\rangle \leq \|\mathbf{x}'\| \|\mathbf{x}\|$$

between $\mathbb{E}$ and its dual $\mathbb{E}'$ immediately yields that $\mathbf{x} \in \mathbb{E}$ can be interpreted, in a canonical way, as an element of the bidual $\mathbb{E}'' = (\mathbb{E}')'$. It is worth to distiguish those spaces for which there is nothing else in $\mathbb{E}''$.

Definition 2.8 A Banach space $\mathbb{E}$ is said to be reflexive, if $\mathbb{E} = \mathbb{E}''$.

The most important example for us of a dual pair is the following.

Theorem 2.2 (Riesz Representation Theorem) *Let $p \in [1, \infty)$, and $\Omega \subset \mathbb{R}^N$, an open subset of finite Lebesgue measure. The dual space of $L^p(\Omega)$ can be identified with $L^q(\Omega)$, $1 = 1/p + 1/q$ (these exponents p and q are called conjugate of each other), through*

$$\Theta : L^q(\Omega) \mapsto L^p(\Omega)', \quad \langle \Theta(g), f\rangle = \int_\Omega f(x)g(x)\, dx \tag{2.14}$$

for $g \in L^q(\Omega)$ and $f \in L^p(\Omega)$.

Proof Through Hölder's inequality, it is elementary to check that the mapping Θ in (2.14) is well-defined, linear, and because

$$\|\Theta(g)\| \leq \|g\|_q,$$

it is also continuous.

We claim that Θ is onto. To see this, take $T \in L^p(\Omega)'$. For a measurable subset $A \subset \Omega$, if $\chi_A(\mathbf{x})$ stands for its characteristic function, given that it belongs to $L^p(\Omega)$, we define a measure m in Ω through formula

$$m(A) = \langle T, \chi_A\rangle. \tag{2.15}$$

It is immediate to check that the set function m is indeed a measure, and, moreover, it is absolutely continuous with respect to the Lebesgue measure because if A is negligible, $\chi_A = 0$ in $L^p(\Omega)$, and so (2.15) yields $m(A) = 0$. The Radon-Nykodim

theorem guarantees that there is some $g \in L^1(\Omega)$ such that

$$\langle T, \chi_A \rangle = \int_\Omega \chi_A(\mathbf{x}) g(\mathbf{x})\, d\mathbf{x}.$$

This identification easily extends to linear combinations of characteristics,[2] and by density to all $f \in L^p(\Omega)$

$$\langle T, f \rangle = \int_\Omega f(\mathbf{x}) g(\mathbf{x})\, d\mathbf{x}, \tag{2.16}$$

provided it is true that $g \in L^q(\Omega)$.

The case $p = \infty$ is clearly correct, since we already know that $g \in L^1(\Omega)$. In particular, (2.16) is valid for uniformly bounded functions f. Suppose $p < \infty$. Put

$$E_n = \{|g| \le n\}, \quad n \in \mathbb{N},$$

and

$$f_n(\mathbf{x}) = \begin{cases} \dfrac{|g(\mathbf{x})|^q}{g(\mathbf{x})}, & \mathbf{x} \in E_n, \\ 0, & else. \end{cases}$$

Since each f_n is uniformly bounded, by our comments above,

$$\langle T, f_n \rangle = \int_\Omega f_n(\mathbf{x}) g(\mathbf{x})\, d\mathbf{x} = \int_{E_n} |g(\mathbf{x})|^q\, d\mathbf{x}.$$

The continuity of T allows to write

$$\int_{E_n} |g(\mathbf{x})|^q\, d\mathbf{x} \le \|T\| \|f_n\|_p = \|T\| \|\chi_{E_n} g\|_q^{q/p},$$

and

$$\|\chi_{E_n} g\|_q \le \|T\|.$$

By the monotone convergence theorem, conclude

$$\|g\|_q \le \|T\|.$$

□

[2] These are called simple functions.

The space $L^\infty(\Omega)$ is very special concerning its dual space. All we can say at this stage is the it contains $L^1(\Omega)$, but it is indeed much larger.

Corollary 2.2 *$L^p(\Omega)$ is reflexive in the finite case $1 < p < \infty$.*

Example 2.10 The precise identification of the dual space of $\mathcal{C}^0(\Omega)$ is technical and delicate, though very important. It requires fundamental tools from Measure Theory. Formally, such dual space can be identified with the space of signed measures with finite total variation on the σ-algebra of the Borel subsets of Ω. It can be generalized for a compact Hausdorff space, or even for a locally compact such space.

2.8 Compactness and Weak Topologies

We remind readers that one of our main initial concerns is that of establishing a suitable compactness principle that may enable us to have a convergent (in some appropriate sense) subsequence from a uniformly bounded sequence. One possibility is to study directly the nature of compact sets of functions in a given Banach space. This is however pretty fruitless, from this perspective, as the topology associated with the norm in a Banach space is usually too fine, and so conditions on compact sets are rather demanding. In the case of L^p-spaces, such compactness criterium is known and important, yet, as we are saying, pretty inoperable in practice, at least from the perspective of our need in this text.

On the other hand, we have already anticipated how, thanks to the Arzelá-Ascoli theorem, bounded sets in $W^{1,p}(J)$ are precompact in $L^\infty(J)$ (if $p > 1$), and sequences converge to functions which indeed remain in $W^{1,p}(J)$. Moreover, we have also identified the kind of convergence that takes place at the level of derivatives

$$\int_J \chi_{(x,y)}(s) f_j'(s)\, ds \to \int_J \chi_{(x,y)}(s) f'(s)\, ds, \tag{2.17}$$

for arbitrary points $x, y \in J$, where, as usual, $\chi_{(x,y)}(s)$ designs the characteristic function of the interval (x, y) in J. It is easy to realize that (2.17) can be extended to simple functions (linear combinations of characteristic functions), and presumably, by some density argument, to more general functions. This short discussion motivates the following fundamental concept.

Definition 2.9 Let $\mathbb{E}$ be a Banach space, with dual $\mathbb{E}'$.

1. We say that a sequence $\{\mathbf{x}_j\} \subset \mathbb{E}$ converges weakly to $\mathbf{x}$, and write $\mathbf{x}_j \rightharpoonup \mathbf{x}$, if for every individual $\mathbf{x}' \in \mathbb{E}'$, we have

$$\langle \mathbf{x}', \mathbf{x}_j \rangle \to \langle \mathbf{x}', \mathbf{x} \rangle$$

as numbers.

2. We say that a sequene $\{\mathbf{x}'_j\} \subset \mathbb{E}'$ converges weakly * to $\mathbf{x}'$, and write $\mathbf{x}'_j \overset{*}{\rightharpoonup} \mathbf{x}'$, if for every individual $\mathbf{x} \in \mathbb{E}$, we have

$$\langle \mathbf{x}'_j, \mathbf{x} \rangle \to \langle \mathbf{x}', \mathbf{x} \rangle$$

as numbers.

This may look, at first sight, as a too general or abstract a concept. It starts to be more promising if we interpret it in our L^p-spaces. In this setting, we would say, after the Riesz representation theorem, that $f_j \rightharpoonup f$ in $L^p(\Omega)$, if

$$\int_\Omega g(\mathbf{x}) f_j(\mathbf{x})\, d\mathbf{x} \to \int_\Omega g(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x}$$

for every $g \in L^q(\Omega)$. If we write this condition in the form

$$\left| \int_\Omega g(\mathbf{x})(f_j(\mathbf{x}) - f(\mathbf{x}))\, d\mathbf{x} \right| \to 0,$$

and compare it to

$$\int_\Omega \left| g(\mathbf{x})(f_j(\mathbf{x}) - f(\mathbf{x})) \right|\, d\mathbf{x} \to 0,$$

we see that the place where the absolute value is put, either outside or inside the integral sign, determines the kind of convergence we are considering (recall Example 2.8). In the first case, we are talking about weak convergence, while in the second about strong or norm convergence. In the case of weak convergence, persistent cancellation phenomenon in the integrals is not excluded; indeed this is the whole point of weak convergence, whereas cancellation is impossible in the second.

In the particular case in which Ω is a finite interval J of $\mathbb{R}$, and g is the characteristic function of a subinterval (x, y), we would find that weak convergence implies

$$\int_x^y f_j(s)\, ds \to \int_x^y f(s)\, ds,$$

exactly the convergence of derivatives that we had anticipated for bounded sequences in the Sobolev space $W^{1,p}(J)$.

Proposition 2.7 *Every bounded sequence in $W^{1,p}(J)$ for $p > 1$, admits a subsequence converging weakly in $W^{1,p}(J)$ and strongly in $L^\infty(J)$, to some function in $W^{1,p}(J)$.*

Proof We have already argued why this statement is correct. Such a bounded sequence $\{f_j\}$ in $W^{1,p}(J)$ is equicontinuous, and hence, for a suitable subsequence which we do not care to relabel, it converges to some f in $L^{\infty}(J)$. Indeed, we showed in Sect. 2.6 that $f \in W^{1,p}(J)$, and that

$$\int_J \chi(x) f_j'(x)\, dx \to \int_J \chi(s) f'(x)\, dx, \tag{2.18}$$

for every characteristic function χ of a subinterval of J. All that remains is to check that (2.18) can be extended from $\chi(x)$ to an arbitrary $g \in L^q(J)$ if q is the conjugate exponent for p. This is done through a standard density argument, since the class of simple functions (linear combinations of characteristic) is dense in any $L^q(\Omega)$. The argument is left as an exercise. □

Possibly, the most important reason to care about weak topologies is the crucial result that follows.

Theorem 2.3 (Banach-Alaouglu-Bourbaki Theorem) *Let $\mathbb{E}$ be a Banach space (it suffices that it be a normed space). The unit ball of the dual $\mathbb{E}'$*

$$\mathbf{B}(\mathbb{E}') = \{\mathbf{x}' \in \mathbb{E}' : \|\mathbf{x}'\| \le 1\}$$

*is compact in the weak * topology.*

Proof Let $\mathbb{L}(\mathbb{E}, \mathbf{K})$ denote the sets of all mappings from $\mathbb{E}$ into $\mathbf{K}$, i.e.

$$\mathbb{L}(\mathbb{E}, \mathbf{K}) = \mathbf{K}^{\mathbb{E}}.$$

It is clear that $\mathbb{E}' \subset \mathbb{L}(\mathbb{E}, \mathbf{K})$, since $\mathbb{E}'$ is the subset of $\mathbb{L}(\mathbb{E}, \mathbf{K})$ of those linear and continuous mappings. It is also easy to realize that

$$\mathbf{B}(\mathbb{E}') \subset \Delta = \prod_{\mathbf{x}\in\mathbb{E}} \Delta_{\mathbf{x}}, \quad \Delta_{\mathbf{x}} = \{\lambda \in \mathbf{K} : |\lambda| \le \|\mathbf{x}\|\},$$

because, for $\mathbf{x}' \in \mathbf{B}(\mathbb{E}')$, it is true that

$$|\langle \mathbf{x}', \mathbf{x}\rangle| \le \|\mathbf{x}\|.$$

The set Δ is compact in the product topology in $\mathbf{K}^{\mathbb{E}}$, because each projection $\Delta_{\mathbf{x}}$, for all $\mathbf{x} \in \mathbb{E}$, is (Tychonoff's theorem). It suffices to check that $\mathbf{B}(\mathbb{E}') \subset \Delta$ is closed (under the weak * topology).

To this end, let $\mathbf{x}'_j \to \mathbf{x}'$ with $\mathbf{x}'_j \in \mathbf{B}(\mathbb{E}')$,[3] with

$$\langle \mathbf{x}'_j, \mathbf{x} \rangle \to \langle \mathbf{x}', \mathbf{x} \rangle \tag{2.19}$$

for each individual $\mathbf{x} \in \mathbb{E}$. It is straightforward to check the linearity of $\mathbf{x}'$ if each $\mathbf{x}'_j$ is. If $\mathbf{x}'_j \in \mathbf{B}(\mathbb{E}')$,

$$|\langle \mathbf{x}'_j, \mathbf{x} \rangle| \le \|\mathbf{x}\|$$

for all j and $\mathbf{x} \in \mathbb{E}$. This implies, through (2.19), that

$$|\langle \mathbf{x}', \mathbf{x} \rangle| \le \|\mathbf{x}\|,$$

i.e. $\mathbf{x}' \in \mathbb{E}'$ (it is continuous), and $\mathbf{x}' \in \mathbf{B}(\mathbb{E}')$. □

We immediately can write a corollary which is the version we will invoke most of the time. If a Banach space is reflexive, then $\mathbb{E} = \mathbb{E}''$, and $\mathbb{E}$ turns out to be the dual of its dual. In this case, the weak star topology in $\mathbb{E}''$ becomes the weak topology in $\mathbb{E}$.

Corollary 2.3 *Let $\mathbb{E}$ be a reflexive, Banach space, and $\{\mathbf{x}_j\}$, a bounded sequence in $\mathbb{E}$. There is always a subsequence converging weakly in $\mathbb{E}$.*

After this corollary the following fundamental fact is at our disposal.

Corollary 2.4 *Let J be a bounded interval in $\mathbb{R}$, and Ω, an open, bounded subset of $\mathbb{R}^N$. Let $p > 1$ be a finite exponent.*

1. *Every bounded sequence in $L^p(\Omega)$ admits a weakly convergent subsequence.*
2. *Every bounded sequence in $W^{1,p}(J)$ admits a subsequence converging weakly in $W^{1,p}(J)$ and strongly in $L^\infty(J)$.*

Example 2.11 We go back to our favorite Example 2.8, but this time we put

$$u'_j(x) = \sin(2^j \pi x), \quad x \in [0, 1].$$

As before, we still have

$$u'_j \rightharpoonup 0 \text{ in } L^2([0, 1]),$$

but we can take

$$u_j(x) = 1 - \frac{1}{2^j \pi} \cos(2^j \pi x),$$

[3] Though we argue here through sequences, in full rigor it should be done through nets. But in this more abstract framework, it is exactly the same.

and it is pretty obvious that

$$u_j \to u \text{ in } L^\infty([0, 1]), \quad u(x) \equiv 1.$$

Example 2.12 Let J be any finite interval in $\mathbb{R}$ of length L, and take $t \in (0, 1)$. There is always a sequence of characteristic functions $\{\chi_j(x)\}$ in J such that

$$\chi_j(x) \rightharpoonup Lt. \tag{2.20}$$

Note that the limit function is a constant. This weak convergence means that one can always find a sequence $\{J_j\}$ of subintervals of J such that

$$\int_{J_j} g(x)\,dx \to Lt \int_J g(x)\,dx$$

for every integrable function g. Take any characteristic function $\chi(x)$ in the unit interval $J_1 = [0, 1]$ such that

$$\int_{J_1} \chi(x)\,dx = t.$$

If $J = (a, a + L)$ then the linear function $x \mapsto Lx + a$ is a bijection between J_1 and J, and hence

$$\int_J \chi(\frac{1}{L}(y - a))\,dy = Lt.$$

If χ is regarded as a 1-periodic function in $\mathbb{R}$, and we put $\chi_j(x) = \chi(2^j x)$, then it is easy to argue that the sequence of functions

$$y \mapsto \chi_j(\frac{1}{L}(y - a))$$

enjoys property (2.20).

All we have said in this and previous sections about $W^{1,p}(J)$ is applicable to Sobolev spaces $W^{1,p}(J; \mathbb{R}^n)$ of paths

$$\mathbf{u}(t) : J \to \mathbb{R}^n, \quad \mathbf{u} = (u_1, u_2, \ldots, u_n),$$

with components $u_i \in W^{1,p}(J)$. We state the main way in which weakly convergent sequences in one-dimensional Sobolev spaces are manipulated in this more general context. It will be invoked later in the book.

Proposition 2.8 *Let* $\{\mathbf{u}_j\} \subset W^{1,p}(J; \mathbb{R}^n)$ *be such that* $\mathbf{u}'_j \rightharpoonup \mathbf{U}$ *in* $L^p(J; \mathbb{R}^n)$*, and* $\mathbf{u}_j(x_0) \to \mathbf{U}_0$ *in* $\mathbb{R}^n$*, for some* $x_0 \in J$*. Then* $\mathbf{u}_j \rightharpoonup \mathbf{u}$ *in* $W^{1,p}(J; \mathbb{R}^n)$ *where*

$$\mathbf{u}(x) = \mathbf{U}_0 + \int_{x_0}^{x} \mathbf{U}(s)\, ds.$$

The proof is left as an exercise.

2.9 Approximation

Another important chapter of Lebesgue and Sobolev spaces is to explicitly specify in what sense smooth functions are dense in these spaces, and how this approximation procedures allow to extend basic results from smooth to non-smooth functions. One initial result in this direction is a classical fact in Measure Theory which in turn makes use of fundamental results in Topology like Urysohn's lemma, or the Tietze extension theorem. Parallel and important results are also Egorov's or Luzin's theorems.

Theorem 2.4 *The set of continuous, real functions with compact support in the real line* $\mathbb{R}$ *is dense in* $L^1(\mathbb{R})$.

This theorem, which is also valid for $\mathbb{R}^N$ and that we take for granted here, will be our basic result upon which build more sophisticated approximation facts.

Lemma 2.4

1. *Let* $f(x) \in L^1(\mathbb{R})$*. There is a sequence* $\{f_j(x)\}$ *of smooth (*$\mathcal{C}^\infty$*), compactly supported functions such that* $f_j \to f$ *in* $L^1(\mathbb{R})$*.*
2. *Let* $f(x) \in W^{1,1}(J)$ *for a finite interval* J*. There is a sequence* $\{f_j(x)\}$ *of smooth (*$\mathcal{C}^\infty$*), functions such that* $f_j \to f$ *in* $W^{1,1}(J)$*.*

The basic technique to achieve this fact makes use of the concept of "mollifier" and convolution that we explain next. Take a smooth $\mathcal{C}^\infty$-function $\rho(z)$, even, non-negative, supported in the symmetric interval $[-1, 1]$ and with

$$\int_{\mathbb{R}} \rho(z)\, dz = \int_{-1}^{1} \rho(z)\, dz = 1. \tag{2.21}$$

One of the most popular choices is

$$\rho(z) = \begin{cases} C \exp \frac{1}{z^2-1}, & |z| \le 1, \\ 0, & |z| \ge 1, \end{cases}$$

for a positive constant to ensure that (2.21) holds. Define

$$f_j(x) = \int_{\mathbb{R}} \rho_j(x-y) f(y)\, dy, \quad \rho_j(z) = j\rho(jz).$$

Note that

$$\int_{\mathbb{R}} \rho_j(z)\, dz = 1, \quad f_j(x) = \int_{\mathbb{R}} \rho_j(y) f(x-y)\, dy.$$

Either of the two formulas to write $f_j(x)$ is called the product of convolution of ρ_j and f, and is written

$$f_j = \rho_j * f.$$

Such a operation between functions enjoys amazing properties. The family $\{\rho_j\}$ is called a mollifier. We claim the following:

1. each f_j is smooth and

$$\frac{d^k}{dx^k} f_j = \frac{d^k}{dx^k} \rho_j * f;$$

2. if the support of f is compact, so is the support of f_j, and all of these supports are contained in a fixed compact;
3. convergence: $f_j \to f$ in $L^1(\Omega)$;

Concerning derivatives, it is easy to realize

$$\frac{1}{h}[f_j(x+h) - f_j(x)] = \int_{\mathbb{R}} \frac{1}{h}[\rho_j(x-y+h) - \rho_j(x-y)] f(y)\, dy,$$

for arbitrary $x, h \in \mathbb{R}$. Since each ρ_j is uniformly bounded (for j fixed), by dominated convergence, as $h \to 0$, we find

$$\lim_{h \to 0} \frac{1}{h}[f_j(x+h) - f_j(x)] = \int_{\mathbb{R}} \rho_j'(x-y) f(y)\, dy.$$

The same reason is valid for the derivatives of any order.

The assertion about the supports is also elementary. If the support of f is contained in an interval $[\alpha, \beta]$, then the support of f_j must be a subset of $[\alpha - 1/j, \beta + 1/j]$. This is easy to realize, because if the distance of x to $[\alpha, \beta]$ is larger than $1/j$, the supports of the two factors in the integral defining f_j are disjoint, and hence the integral vanishes.

Finally, for the convergence we follow a very classical and productive technique, based on Theorem 2.4, consisting in proving the desired convergence for continuous functions with compact support, and then take advantage of the density fact of

Theorem 2.4. Assume then that $g(x)$ is a real function, which is continuous and with compact support. In particular, g is uniformly continuous. If we examine the difference (bear in mind that the integral of every ρ_j is unity)

$$g_j(x) - g(x) = \int_{\mathbb{R}} \rho_j(y)(g(x-y) - g(x))\,dy,$$

because the support of ρ_j is the interval $[-1/j, 1/j]$, and the uniform continuity of g, given $\epsilon > 0$, j_0 can be found so that

$$|g(x-y) - g(x)| \le \epsilon, \quad j \ge j_0$$

for every y with $|y| \le 1/j$, and all $x \in \mathbb{R}$. Thus

$$|g_j(x) - g(x)| \le \epsilon$$

for all $x \in \mathbb{R}$ if $j \ge j_0$. If we bring to mind, because of our second claim above about the supports, that the supports of all g_j's and g are contained in a fixed compact, then the previous uniform estimate implies that $g_j \to g$ in $L^1(\mathbb{R})$.

Proof of Lemma 2.4 In the context of the preceding discussion, consider $f \in L^1(\mathbb{R})$ arbitrary, and let $\epsilon > 0$ be given. By Theorem 2.4, select a continuous function g with compact support such that

$$\|f - g\|_{L^1(\mathbb{R})} \le \epsilon.$$

We write

$$\|f_j - f\|_{L^1(\mathbb{R})} \le \|f_j - g_j\|_{L^1(\mathbb{R})} + \|g_j - g\|_{L^1(\mathbb{R})} + \|g - f\|_{L^1(\mathbb{R})}.$$

The last term is smaller than ϵ, while the second one can be made also that small, provided j is taken sufficiently large. Concerning the first one, we expand

$$\int_{\mathbb{R}} |f_j(x) - g_j(x)|\,dx \le \int_{\mathbb{R}} \int_{\mathbb{R}} \rho_j(x-y)|f(y) - g(y)|\,dy\,dx.$$

A change of the order of integration[4] carries us directly to

$$\|f_j - g_j\|_{L^1(\mathbb{R})} \le \|f - g\|_{L^1(\mathbb{R})} \le \epsilon,$$

as well. The first item in the lemma is proved.

[4] We leave to the interested readers to check the technical conditions required for the validity of this fact.

Concerning the second, it is achieved by integration. If we apply the first to the $L^1(\mathbb{R})$-function $f'(x)\chi_J(x)$, where $\chi_J(x)$ is the characteristic function of J, we would find a sequence of smooth functions g_j with $f'-g_j$ tending to zero in $L^1(\mathbb{R})$. We put

$$f_j(x) = \int_{-\infty}^{x} g_j(y)\,dy.$$

Then $f_j \in W^{1,1}(\mathbb{R})$,

$$f_j'(x) = g_j(x), \quad \|f_j' - f'\|_{L^1(\mathbb{R})} \to 0,$$

and, as usual, the convergence $f_j - f$ to zero takes place, through integration of its derivative, in $L^\infty(\mathbb{R})$. The restriction of all functions to J yields the result. □

This approximation fact, which is just one example of a whole family of such results, permits us to generalize results, that are very classical when smoothness assumptions are guaranteed, to the context of functions in Sobolev spaces. Some times there is no other way of showing the validity of such results. The following is the product rule and the generalization of the integration-by-parts formula for functions in $W^{1,1}(\mathbb{R})$.

Proposition 2.9 *Let $J = [\alpha, \beta]$ be a finite interval. Assume f and g are functions belonging to $W^{1,1}(J)$.*

1. *The product fg also belongs to the same space, and*

$$(fg)' = f'g + fg'.$$

2. *We have*

$$\int_J f'(x)g(x)\,dx = -\int_J f(x)g'(x)\,dx + f(\beta)g(\beta) - f(\alpha)g(\alpha).$$

In particular, if one of the two functions, either f or g, belongs to $W_0^{1,1}(J)$ (recall Definition 2.6), then

$$\int_J f'(x)g(x)\,dx = -\int_J f(x)g'(x)\,dx. \tag{2.22}$$

Proof Suppose first that the factor g is smooth, and so belonging to $W^{1,1}(J)$. The weak derivative of f is such that

$$\int_J f'(x)g(x)\phi(x)\,dx = -\int_J f(x)(g(x)\phi(x))'\,dx$$

for every smooth function ϕ with compact support in J, or vanishing at both end-points α and β. Note that these facts at the end-points of J are also correct for the product $g\phi$ if g is smooth. But we know that the product rule is valid when the factors are smooth and we are dealing with classical derivatives. Then

$$\int_J f'(x)g(x)\phi(x)\,dx = -\int_J [f(x)g'(x)\phi(x) + f(x)g(x)\phi'(x)]\,dx,$$

which can be reorganized in the form

$$\int_J f(x)g(x)\phi'(x)\,dx = -\int_J (f'(x)g(x) + f(x)g'(x))\phi(x)\,dx.$$

The arbitrariness of ϕ indicates that the product fg admits as a weak derivative $f'g + fg'$ which is a function in $L^1(J)$.

If g is a general function in $W^{1,1}(J)$, we proceed by approximation. Thanks to our previous lemma, we can find a sequence g_j of smooth functions converging to g in $W^{1,1}(J)$. For a smooth, test function $\phi(x)$, we write, by our previous step,

$$\begin{aligned}
\int_J f(x)g(x)\phi'(x)\,dx &= \lim_{j\to\infty}\int_J f(x)g_j(x)\phi'(x)\,dx \\
&= -\lim_{j\to\infty}\int_J (f'(x)g_j(x) + f(x)g_j'(x))\phi(x)\,dx \\
&= -\int_J (f'(x)g(x) + f(x)g'(x))\phi(x)\,dx.
\end{aligned}$$

The arbitrariness of ϕ implies our result. The formula of integration by parts is a straightforward consequence of the product rule. □

2.10 Completion of Spaces of Smooth Functions with Respect to Integral Norms

We use here Theorem 2.1 to build Lebesgue spaces and Sobolev spaces in dimension one. The process for Sobolev spaces in higher dimension is similar, as we will check later in the book.

Consider the vector space $\mathcal{C}^\infty(\Omega)$ of smooth functions in an open set $\Omega \subset \mathbb{R}^N$. As in Sect. 2.4, we define the pth- norm of functions in $\mathcal{C}^\infty(\Omega)$ for $1 \le p \le \infty$. Young's and Hölder's inequalities imply that the pth-norm is indeed a norm in this space of smooth functions, and hence, by Theorem 2.1, we can consider its completion $\mathcal{L}^p(\Omega)$ with respect to this norm. The only point that deserves some special comment is the fact that the pth-norm is just a seminorm, not a norm, in this

completion, and it becomes again a norm on the quotient space

$$L^p(\Omega) = \mathcal{L}^p(\Omega)/\mathcal{N}^p$$

where $\mathcal{N}^p$ is the class of functions that vanish except for a negligible set. The concepts of subnorm and seminorm will be introduced and used in Chap. 4.

To be honest, this process yields a Banach subspace of the true $L^p(\Omega)$, in which smooth functions are dense (with respect to the pth-norm). To be sure that such a subspace is in fact the full Banach space $L^p(\Omega)$ requires to check that in these spaces smooth functions are also dense. But this demands some extra work focused on approximation, which is interesting but technical, as we have seen in the preceding section.

Concerning Sobolev spaces, the process is similar but the completion procedure in Theorem 2.1 is performed with respect to the Sobolev norm in Proposition 2.3. The resulting space is a Banach subspace (again after making the quotient over the class of a.e. null functions) of the true Sobolev space $W^{1,p}(J)$. They are the same spaces, but this asks for checking that the space of smooth functions is dense in $W^{1,p}(J)$, which once more is an important approximation procedure.

The process in higher dimension is not more difficult, though this important approximation fact of Sobolev functions by smooth functions is more involved. We will recall these ideas later in Chap. 7 when we come to studying higher-dimensional Sobolev spaces, and stress how, from the viewpoint of variational problems, one can work directly with these completions of smooth functions under the appropriate norms.

2.11 Hilbert Spaces

There is a very special class of Banach spaces which share with finite-dimensional Euclidean spaces one very fundamental feature: an inner product.

Definition 2.10 Let $\mathbb{H}$ be a vector space.

1. An inner or scalar product in $\mathbb{H}$ is a bilinear, symmetric, positive definite, non-degenerate form

$$\langle\cdot,\cdot\rangle : \mathbb{H} \times \mathbb{H} \to \mathbb{R}.$$

The norm in $\mathbb{H}$ is induced by the inner product through the formula

$$\|\mathbf{x}\|^2 = \langle\mathbf{x}, \mathbf{x}\rangle \geq 0. \tag{2.23}$$

2. A Hilbert space $\mathbb{H}$ is a vector space endowed with an inner product which is a Banach space under the corresponding norm (2.23).

The properties required for a function

$$\langle \cdot, \cdot \rangle : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$$

to be an inner product explicitly are:

1. symmetry: for every $\mathbf{x}, \mathbf{y} \in \mathbb{H}$,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle;$$

2. linearity on each entry: for every fixed $\mathbf{y} \in \mathbb{H}$,

$$\langle \mathbf{y}, \cdot \rangle : \mathbb{H} \to \mathbb{R}, \quad \langle \cdot, \mathbf{y} \rangle : \mathbb{H} \to \mathbb{R},$$

   are linear functionals;
3. positive definiteness:

$$\langle \mathbf{x}, \mathbf{x} \rangle \geq 0, \quad \langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}.$$

The triangle inequality $N3$ for the norm coming form an inner product is a consequence of the Cauchy-Schwarz inequality that is correct for every scalar product

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|.$$

The model Hilbert space is $L^2(\Omega)$ for an open subset $\Omega \subset \mathbb{R}^N$. It is, by far, the most important Hilbert space in Applied Analysis and Mathematical Physics. Similarly, if J is an interval in $\mathbb{R}$, the Sobolev space $W^{1,2}(J)$ is also a Hilbert space.

Proposition 2.10 *The mapping*

$$\langle f, g \rangle = \int_\Omega f(\boldsymbol{x}) g(\boldsymbol{x})\, d\boldsymbol{x}$$

defines an inner product on the set of measurable functions defined in Ω*, with norm*

$$\|f\|_2^2 = \int_\Omega f(\boldsymbol{x})^2\, d\boldsymbol{x}.$$

The space $L^2(\Omega)$ *of square-integrable, measurable functions is a Hilbert space under this inner product. In a similar way, the map*

$$\langle f, g \rangle = \int_J [f(x)g(x) + f'(x)g'(x)]\, dx$$

determines an inner product in $W^{1,2}(J)$. This space is typically referred to as $H^1(J)$.

All that the proof of this proposition requires, knowing already that $L^2(\Omega)$ and $W^{1,2}(J)$ are Banach spaces under their respective 2-norms, is to check that, in both cases, the formula determining the inner product is symmetric, bi-linear, and positive definite. This is elementary.

The fact that a Banach space can be endowed with a inner product under which it becomes a Hilbert space may look like something accidental, or simply convenient. But it is not so, as an inner product allows for fundamental and profound facts:

1. In a Hilbert space, one can talk about orthogonal projections onto subspaces, and even onto convex, closed subsets.
2. Orthogonality also has important consequences as one can consider orthonormal bases.
3. The inner product permits to identify a Hilbert space with its own dual whenever convenient. In this sense, we can say that $\mathbb{H} = \mathbb{H}'$.
4. The extension of many facts from multivariate Calculus to infinite dimension is possible thanks to the inner product.

We treat these four important issues successively.

2.11.1 Orthogonal Projection

We start with one of the most remarkable and fundamental tools. Recall that a convex set $\mathbb{K}$ of a vector space $\mathbb{H}$ is such that convex combinations

$$t\mathbf{x}_1 + (1-t)\mathbf{x}_0 \in \mathbb{K}$$

whenever $\mathbf{x}_1, \mathbf{x}_0 \in \mathbb{K}$ and $t \in [0, 1]$.

Proposition 2.11 (Projection Onto a Convex Set) *Let $\mathbb{K} \subset \mathbb{H}$ be a closed and convex subset of a Hilbert space $\mathbb{H}$. For every $\mathbf{x} \in \mathbb{H}$, there is a unique $\mathbf{y} \in \mathbb{K}$ such that*

$$\|\mathbf{x} - \mathbf{y}\| = \min_{\mathbf{z} \in \mathbb{K}} \|\mathbf{x} - \mathbf{z}\|.$$

Moreover, $\mathbf{y}$ is characterized by the condition

$$(\mathbf{y} - \mathbf{x}) \cdot (\mathbf{z} - \mathbf{y}) \geq 0$$

for every $\mathbf{z} \in \mathbb{K}$.

Proof Let $\mathbf{y}_j \in \mathbb{K}$ be a minimizing sequence for the problem

$$m = \inf_{\mathbf{z}\in\mathbb{K}} \|\mathbf{z} - \mathbf{x}\|^2, \quad \|\mathbf{y}_j - \mathbf{x}\|^2 \searrow m.$$

From the usual parallelogram law, which is also valid in a general Hilbert space,

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2, \quad \mathbf{x}, \mathbf{y} \in \mathbb{H},$$

we find that

$$\|\mathbf{y}_j - \mathbf{y}_k\|^2 = 2\|\mathbf{y}_j - \mathbf{x}\|^2 + 2\|\mathbf{y}_k - \mathbf{x}\|^2 - 4\|\mathbf{x} - (1/2)(\mathbf{y}_j + \mathbf{y}_k)\|^2.$$

Since $\mathbb{K}$ is convex, the middle point

$$\frac{1}{2}(\mathbf{y}_j + \mathbf{y}_k)$$

belongs to $\mathbb{K}$ as well. Hence, by definition of m,

$$\|\mathbf{y}_j - \mathbf{y}_k\|^2 \le 2\|\mathbf{y}_j - \mathbf{x}\|^2 + 2\|\mathbf{y}_k - \mathbf{x}\|^2 - 4m.$$

The right-hand side converges to zero as $j, k \to \infty$ because $\{\mathbf{y}_j\}$ is minimizing. This implies that $\{\mathbf{y}_j\}$ is a Cauchy sequence, and so, it converges to some $\mathbf{y}$ in $\mathbb{K}$ (because $\mathbb{K}$ is closed), which is the minimizer. The uniqueness follows from the strict convexity of the norm in a Hilbert space (or again from the parallelogram rule).

For arbitrary $\mathbf{z} \in \mathbb{K}$, consider the convex combination

$$t\mathbf{z} + (1 - t)\mathbf{y} \in \mathbb{K}, \quad t \in [0, 1].$$

The function

$$\phi(t) = \|\mathbf{x} - t\mathbf{z} - (1 - t)\mathbf{y}\|^2, \quad t \in [0, 1]$$

attains its minimum at $t = 0$, and so $\phi'(0) \ge 0$. But

$$\phi'(0) = 2(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{y} - \mathbf{z}).$$

Conversely, if for every $\mathbf{z} \in \mathbb{K}$, we have that

$$0 \le (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{y} - \mathbf{z}) = -\|\mathbf{x} - \mathbf{y}\|^2 + (\mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{z}),$$

then, by the Cauchy-Schwarz inequality,

$$\|\mathbf{x} - \mathbf{y}\|^2 \le (\mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{z}) \le \|\mathbf{x} - \mathbf{y}\| \, \|\mathbf{x} - \mathbf{z}\|$$

for all such $\mathbf{z}$. This implies that

$$\|\mathbf{x} - \mathbf{y}\| = \min_{\mathbf{z} \in \mathbb{K}} \|\mathbf{x} - \mathbf{z}\|,$$

as desired. □

This lemma permits to consider the map

$$\pi_{\mathbb{K}} : \mathbb{H} \to \mathbb{K},$$

called the projection onto $\mathbb{K}$, defined precisely by putting

$$\pi_{\mathbb{K}}(\mathbf{x}) = \mathbf{y},$$

the unique vector $\mathbf{y}$ in the statement of the last proposition. The projection $\pi_{\mathbb{K}}$ is a continuous map. In fact, by the characterization of the projection in the previous proposition, we have

$$(\pi_{\mathbb{K}}\mathbf{x}_1 - \mathbf{x}_1) \cdot (\pi_{\mathbb{K}}\mathbf{x}_2 - \pi_{\mathbb{K}}\mathbf{x}_1) \geq 0,$$
$$(\pi_{\mathbb{K}}\mathbf{x}_2 - \mathbf{x}_2) \cdot (\pi_{\mathbb{K}}\mathbf{x}_1 - \pi_{\mathbb{K}}\mathbf{x}_2) \geq 0.$$

By adding these two inequalities, we can write

$$\begin{aligned} \|\pi_{\mathbb{K}}\mathbf{x}_1 - \pi_{\mathbb{K}}\mathbf{x}_2\|^2 =& (\pi_{\mathbb{K}}\mathbf{x}_1 - \pi_{\mathbb{K}}\mathbf{x}_2) \cdot (\pi_{\mathbb{K}}\mathbf{x}_1 - \pi_{\mathbb{K}}\mathbf{x}_2) \\ &\leq (\mathbf{x}_1 - \mathbf{x}_2) \cdot (\pi_{\mathbb{K}}\mathbf{x}_1 - \pi_{\mathbb{K}}\mathbf{x}_2), \end{aligned}$$

and, again by the Cauchy-Schwarz inequality,

$$\|\pi_{\mathbb{K}}\mathbf{x}_1 - \pi_{\mathbb{K}}\mathbf{x}_2\| \leq \|\mathbf{x}_1 - \mathbf{x}_2\|.$$

The very particular case in which $\mathbb{K}$ is a closed subspace is especially relevant.

Corollary 2.5 *If $\mathbb{K}$ is a closed, subspace of a Hilbert space $\mathbb{H}$, the orthogonal projection $\pi_{\mathbb{K}}$ is a linear, continuous operator characterized by the condition*

$$\langle \mathbf{x} - \pi_{\mathbb{K}}\mathbf{x}, \mathbf{y} \rangle = 0$$

for every $\mathbf{y} \in \mathbb{K}$.

2.11.2 Orthogonality

The existence of an inner product leads to the fundamental geometrical concept of orthogonality.

Definition 2.11 A sequence of vectors $\{\mathbf{x}_j\}$ of a Hilbert space $\mathbb{H}$ is said to be a orthonormal basis if

$$\|\mathbf{x}_j\| = 1 \text{ for all } j, \quad \langle \mathbf{x}_j, \mathbf{x}_k \rangle = 0 \text{ for all pairs } j \neq k,$$

and the subspace spanned by it is dense in $\mathbb{H}$.

Hilbert spaces that admit a orthonormal basis need to be separable, meaning by that the following.

Definition 2.12 A Banach space $\mathbb{E}$ is said to be separable if there is a countable subset which is dense.

The fundamental facts about orthogonality in Hilbert spaces can be specified through the following two results.

Proposition 2.12 *Every separable Hilbert space admits orthonormal bases.*

Proof Let $\{\overline{\mathbf{x}}_j\}$ be dense in the Hilbert space $\mathbb{H}$, and let $\mathbb{H}_j$ be the subspace spanned by $\{\overline{\mathbf{x}}_k\}_{k\leq j}$, so that $\{\mathbb{H}_j\}$ is a non-decreasing sequence of finite-dimensional subspaces whose union is dense in $\mathbb{H}$. By using the Gram-Schmidt orthonormalization process, we can produce an increasing sequence of orthonormal vectors $\{\mathbf{x}_k\}$ that spanned the successive subspaces $\mathbb{H}_j$. This full collection of vectors $\{\mathbf{x}_k\}$ is a orthonormal basis of $\mathbb{H}$ according to Definition 2.11. □

The following is also a fundamental feature of Hilbert spaces.

Proposition 2.13 *Let $\mathbb{H}$ be a separable, Hilbert space, and $\{\mathbf{x}_j\}$, a orthonormal basis.*

1. *For every $\mathbf{x} \in \mathbb{H}$,*

$$\mathbf{x} = \sum_{j=1}^{\infty} \langle \mathbf{x}, \mathbf{x}_j \rangle \mathbf{x}_j, \quad \|\mathbf{x}\|^2 = \sum_{j=1}^{\infty} |\langle \mathbf{x}, \mathbf{x}_j \rangle|^2.$$

2. *If the sequence of numbers $\{a_j\}$ is square-summable*

$$\sum_{j=1}^{\infty} a_j^2 < \infty,$$

then

$$\mathbf{x} = \sum_{j=1}^{\infty} a_j \mathbf{x}_j \in \mathbb{H}, \quad \|\mathbf{x}\|^2 = \sum_{j=1}^{\infty} a_j^2, \quad a_j = \langle \mathbf{x}, \mathbf{x}_j \rangle.$$

Proof Let $\mathbf{x} \in \mathbb{H}$ be given, put $\mathbb{H}_j$ for the subspace spanned by $\{\mathbf{x}_k\}_{k \le j}$, and

$$\mathbf{x}^{(j)} = \sum_{k=1}^{j} \langle \mathbf{x}, \mathbf{x}_k \rangle \mathbf{x}_k.$$

If $\pi_{\mathbb{H}_j}$ is the orthogonal projection onto $\mathbb{H}_j$, we claim that

$$\mathbf{x}^{(j)} = \pi_{\mathbb{H}_j} \mathbf{x}. \tag{2.24}$$

To check this, consider the optimization problem

$$\text{Minimize in } \mathbf{z} = (z_k) \in \mathbb{R}^j : \quad \|\mathbf{x} - \sum_{k=1}^{j} z_k \mathbf{x}_k\|^2.$$

We can actually write, taking into account the orthogonality of $\{\mathbf{x}_k\}$ and the properties of the inner product,

$$\begin{aligned} \|\mathbf{x} - \sum_{k=1}^{j} z_k \mathbf{x}_k\|^2 =& \langle \mathbf{x} - \sum_{k=1}^{j} z_k \mathbf{x}_k, \mathbf{x} - \sum_{k=1}^{j} z_k \mathbf{x}_k \rangle \\ =& \|\mathbf{x}\|^2 - 2 \sum z_k \langle \mathbf{x}, \mathbf{x}_k \rangle + \sum z_k^2. \end{aligned}$$

It is an elementary Vector Calculus exercise to find that the optimal solution $\overline{\mathbf{z}}$ exactly corresponds to

$$\overline{z}_k = \langle \mathbf{x}, \mathbf{x}_k \rangle.$$

This means that indeed (2.24) holds. In particular, thanks to Corollary 2.5, we can conclude that

$$\langle \mathbf{x} - \mathbf{x}^{(j)}, \mathbf{x}^{(j)} \rangle = \langle \mathbf{x} - \pi_{\mathbb{H}_j} \mathbf{x}, \pi_{\mathbb{H}_j} \mathbf{x} \rangle = 0. \tag{2.25}$$

Bearing in mind this fact, we further realize that

$$0 \le \langle \mathbf{x} - \mathbf{x}^{(j)}, \mathbf{x} - \mathbf{x}^{(j)} \rangle = \|\mathbf{x}\|^2 - \langle \mathbf{x}, \mathbf{x}^{(j)} \rangle,$$

that is to say

$$\sum_{k=1}^{j} |\langle \mathbf{x}, \mathbf{x}_k \rangle|^2 \leq \|\mathbf{x}\|^2,$$

for arbitrary j. This implies that the series of non-negative numbers

$$\sum_{k=1}^{j} |\langle \mathbf{x}, \mathbf{x}_k \rangle|^2$$

is convergent. But, again bearing in mind the orghogonality of $\{\mathbf{x}_j\}$, we can also conclude for $k < j$, that

$$\|\mathbf{x}^{(j)} - \mathbf{x}^{(k)}\|^2 = \sum_{l=k+1}^{j} |\langle \mathbf{x}, \mathbf{x}_l \rangle|^2.$$

Hence $\{\mathbf{x}^{(j)}\}$ is a Cauchy sequence, and it converges to some $\overline{\mathbf{x}}$. It remains to show that in fact $\overline{\mathbf{x}} = \mathbf{x}$.

The condition

$$\overline{\mathbf{x}} = \sum_{k=1}^{\infty} \langle \mathbf{x}, \mathbf{x}_k \rangle \mathbf{x}_k,$$

implies, just as we did with $\mathbf{x}$, that

$$\pi_{\mathbb{H}_j} \overline{\mathbf{x}} = \sum_{k=1}^{j} \langle \mathbf{x}, \mathbf{x}_k \rangle \mathbf{x}_k = \pi_{\mathbb{H}_j} \mathbf{x}.$$

Then

$$\pi_{\mathbb{H}_j} (\overline{\mathbf{x}} - \mathbf{x}) = \mathbf{0},$$

for all j, and hence the difference $\overline{\mathbf{x}} - \mathbf{x}$ is a vector which is orthogonal to the full basis $\{\mathbf{x}_j\}$. This implies that indeed $\overline{\mathbf{x}} - \mathbf{x}$ must be the vanishing vector, for otherwise the subspace $\overline{\mathbb{H}}$ spanned by the full basis could not be dense in $\mathbb{H}$, as it would accept a non-vanishing orthogonal vector.

The second part of the statement requires exactly the same ideas. $\square$

It is quite instructive to look at some explicit cases of orthonormal bases. The first one is mandatory.

Example 2.13 Consider the Hilbert space

$$\ell^2 \equiv \ell^2(\mathbb{R}) = \{\mathbf{x} = (x_k) : \sum_k x_k^2 < \infty\},$$

with inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_k x_k y_k.$$

If we let $\mathbf{e}_k \in \ell^2$ be the trivial sequence except for 1 in the k-th place, then it is evident that the countable collection $\{\mathbf{e}_k : k \in \mathbb{N}\}$, the canonical basis, is an orthonormal basis of ℓ^2.

Notice how Proposition 2.13 establishes that every separable real Hilbert space is isomorphic to ℓ^2.

The most important and popular example is possibly that of the trigonometric basis in $L^2(-\pi, \pi)$.

Example 2.14 Consider the family of trigonometric functions

$$\{1\} \cup \{\cos(kx), \sin(kx)\}_{k \in \mathbb{N}}.$$

It is elementary to check that they belong to $L^2(-\pi, \pi)$: by the double-angle formula, we know that

$$\int_{-\pi}^{\pi} \cos^2(kx)\, dx = \frac{1}{2} \int_{-\pi}^{\pi} [1 + \cos(2kx)]\, dx = \pi,$$

$$\int_{-\pi}^{\pi} \sin^2(kx)\, dx = \frac{1}{2} \int_{-\pi}^{\pi} [1 - \cos(2kx)]\, dx = \pi,$$

for all $k \in \mathbb{N}$. Moreover, the elementary trigonometric formulas

$$\cos(\alpha \pm \beta) = \cos\alpha \, \cos\beta \mp \sin\alpha \, \sin\beta,$$

$$\sin(\alpha \pm \beta) = \sin\alpha \, \cos\beta \pm \cos\alpha \, \sin\beta,$$

suitably utilized, lead to the identities

$$\int_{-\pi}^{\pi} \cos(kx)\cos(jx)\, dx = 0, \quad k \neq j,$$

$$\int_{-\pi}^{\pi} \sin(kx)\sin(jx)\, dx = 0, \quad k \neq j,$$

$$\int_{-\pi}^{\pi} \cos(kx)\sin(jx)\, dx = 0, \quad k \neq j.$$

All these calculations prove that the family of functions

$$\mathcal{F} = \{\frac{1}{\sqrt{2\pi}}\} \cup \{\frac{1}{\sqrt{\pi}} \cos(kx), \frac{1}{\sqrt{\pi}} \sin(kx)\} \tag{2.26}$$

is orthonormal. In fact, if we let

$$y_k(x) = \cos(kx), \quad z_k(x) = \sin(kx) = -\frac{1}{k} y_k'(x), \quad k \in \mathbb{N},$$

it is easy to realize that

$$y_k'' + k^2 y_k = 0, \quad z_k'' + k^2 z_k = 0. \tag{2.27}$$

But then, relying on two integration by parts for which end-point contributions drop out,

$$\begin{aligned} k^2 \int_{-\pi}^{\pi} y_k(x) y_j(x)\, dx &= - \int_{-\pi}^{\pi} y_k''(x) y_j(x)\, dx \\ &= \int_{-\pi}^{\pi} y_k'(x) y_j'(x)\, dx \\ &= - \int_{-\pi}^{\pi} y_k(x) y_j''(x)\, dx \\ &= j^2 \int_{-\pi}^{\pi} y_k(x) y_j(x)\, dx. \end{aligned}$$

If $k \neq j$, we conclude that $y_k(x)$ and $y_j(x)$ are orthogonal in $L^2(-\pi, \pi)$, without the need to compute the definite integrals explicitly. Similar manipulations lead to the other two orthogonality relations.

Even more is true. The Fourier family of functions (2.26) is indeed a basis for $L^2(-\pi, \pi)$. This can be checked directly by showing that Fourier partial sums of the form

$$a_0 + \sum_{k=1}^{N} [a_k \cos(kx) + b_k \sin(kx)]$$

with

$$a_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \cos(kx)\, dx, \quad a_0 = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x)\, dx,$$

$$b_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \sin(kx)\, dx,$$

converge, in $L^2(-\pi, \pi)$, to an arbitrary function $f \in L^2(-\pi, \pi)$ as $N \to \infty$, though it requires a good deal of fine work. But similar trigonometric families of functions are shown to be bases of L^2-spaces as a remarkable consequence of (2.27). We will later explore this in Theorem 6.4.

There are other fundamental examples of orthonormal basis that are introduced in the exercises below.

2.11.3 The Dual of a Hilbert Space

The inner product in a Hilbert space $\mathbb{H}$ also has profound consequences for the dual $\mathbb{H}'$.

Proposition 2.14 (Fréchet-Riesz Representation Theorem) *For every* $\mathbf{x}' \in \mathbb{H}'$ *there is a unique* $\mathbf{x} \in \mathbb{H}$ *such that*

$$\langle \mathbf{x}', \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle, \quad \mathbf{y} \in \mathbb{H}.$$

We are using brackets here for two different things at first sight: the left-hand side corresponds to the duality between $\mathbb{H}'$ and $\mathbb{H}$, while in the right-hand side, it is the inner product in $\mathbb{H}$.

Proof Let $\mathbf{x}' \in \mathbb{H}'$ be arbitrary, and put

$$\mathbb{H}_0 = \{\mathbf{y} \in \mathbb{H} : \langle \mathbf{x}', \mathbf{y} \rangle = 0\} \subset \mathbb{H},$$

a closed subspace of $\mathbb{H}$. If $\mathbb{H}_0 = \mathbb{H}$, then $\mathbf{x}' = \mathbf{0}$, and we can, obviously, take $\mathbf{x} = \mathbf{0}$ as well. Assume then that $\mathbb{H}_0$ is not the full $\mathbb{H}$.

Take $\mathbf{x}_0 \in \mathbb{H} \setminus \mathbb{H}_0$, a non-vanishing vector, and put

$$\mathbf{X}_0 = \frac{\mathbf{x}_0 - \pi_{\mathbb{H}_0}\mathbf{x}_0}{\|\mathbf{x}_0 - \pi_{\mathbb{H}_0}\mathbf{x}_0\|}.$$

The norm in the denominator cannot vanish precisely because $\mathbf{x}_0 \notin \mathbb{H}_0$. For $\mathbf{y} \in \mathbb{H}$ arbitrary, the combination

$$\mathbf{y} - \frac{\langle \mathbf{x}', \mathbf{y} \rangle}{\langle \mathbf{x}', \mathbf{X}_0 \rangle}\mathbf{X}_0 \tag{2.28}$$

belongs to $\mathbb{H}_0$ because $\mathbf{x}'$ applied to it, vanishes. The number in the denominator does not vanish because, once again $\mathbf{x}_0 \notin \mathbb{H}_0$. By Corollary 2.5, applied to $\mathbf{x} = \mathbf{x}_0$, $\mathbb{K} = \mathbb{H}_0$, and $\mathbf{y}$, the combination in (2.28), we conclude that

$$\langle \mathbf{X}_0, \mathbf{y} - \frac{\langle \mathbf{x}', \mathbf{y} \rangle}{\langle \mathbf{x}', \mathbf{X}_0 \rangle}\mathbf{X}_0 \rangle = 0,$$

as well for all $\mathbf{y} \in \mathbb{H}$, and this identity lets us see that it suffices to take

$$\mathbf{x} = \langle \mathbf{x}', \mathbf{X}_0 \rangle \mathbf{X}_0.$$

Recall that $\mathbf{X}_0$ is unitary. □

This theorem clearly establishes that a Hilbert space can always be identified, in a canonical way through the inner product, with its own dual. In particular, every Hilbert space is reflexive.

2.11.4 Basic Calculus in a Hilbert Space

The structure of a Hilbert space through its inner product allows for many similar facts as in finite dimensional spaces. The following definition refers to a functional $I : \mathbb{H} \to \mathbb{R}$ where $\mathbb{H}$ is a Hilbert space.

Definition 2.13

1. Such a functional I is said to be Gateaux-differentiable if every section, for arbitrary $\mathbf{x}, \mathbf{y} \in \mathbb{H}$,
$$\epsilon \mapsto I(\mathbf{x} + \epsilon \mathbf{y})$$
is differentiable as a function of the single variable ϵ.
2. I is differentiable if it is Gateaux-differentiable, and the operation
$$(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{T}(\mathbf{x}, \mathbf{y}) = \left. \frac{d}{d\epsilon} I(\mathbf{x} + \epsilon \mathbf{y}) \right|_{\epsilon=0}$$
is continuous in $\mathbf{x}$, and linear in $\mathbf{y}$, in such a way that we can write
$$\mathbf{T}(\mathbf{x}, \mathbf{y}) = \langle \tilde{\mathbf{T}}\mathbf{x}, \mathbf{y} \rangle,$$
for a certain continuous $\tilde{\mathbf{T}} : \mathbb{H} \to \mathbb{H}$. We put
$$I'(\mathbf{x}) = \tilde{\mathbf{T}}\mathbf{x},$$
and refer to it as the derivative of I. Note that $I' : \mathbb{H} \to \mathbb{H}$ is a continuous, in general non-linear, operation.
3. A certain element $\mathbf{x} \in \mathbb{H}$ is said to be critical for a differentiable functional I, if $I'(\mathbf{x}) = \mathbf{0}$.

4. The functional I as above is Fréchet differentiable if it is differentiable and, for every $\mathbf{x} \in \mathbb{H}$,

$$\frac{1}{\|\mathbf{y}\|}\|I(\mathbf{x}+\mathbf{y}) - I(\mathbf{x}) - \langle I'(\mathbf{x}), \mathbf{y}\rangle\| \to 0 \text{ as } \mathbf{y} \to \mathbf{0}. \tag{2.29}$$

The concepts of differentiability and Fréchet differentiability can be localized at individual vectors $\mathbf{x}$, but since most of the time they are used in a global way, we do not consider those local definitions. On the other hand, we are incorporating the continuity of the derivative into the differentiability. Usually this is not done so in other sources. Again, in most of the important examples and applications, the continuity of the derivative is required, and thus we have judged more transparent to proceed in this way.

A parallelism with functions of several variables can be clearly established. If

$$f(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}$$

is a function, then Gateaux-differentiability amounts to the possibility of computing all directional derivatives; differentiability means linearity on those directional derivatives, and continuity of partial derivatives; and Fréchet-differentiability amounts to plain differentiability. The well-known criterium in finite dimension that the continuity of partial derivatives implies differentiability also has a counterpart in infinite dimension.

Lemma 2.5 *A functional $I : \mathbb{H} \to \mathbb{R}$ defined on a Hilbert space is differentiable if and only if it is Fréchet-differentiable.*

Proof Suppose I is differentiable with continuous derivative $I' : \mathbb{H} \to \mathbb{H}$. Let $\mathbf{x}, \mathbf{y} \in \mathbb{H}$, and consider the real function

$$g : \mathbb{R} \to \mathbb{R}, \quad g(s) = I(\mathbf{x} + s\mathbf{y}).$$

We know that this function is continuously differentiable and besides, because I is differentiable,

$$g'(s) = \langle I'(\mathbf{x} + s\mathbf{y}), \mathbf{y}\rangle.$$

By the mean-value theorem (for real functions)

$$g(1) - g(0) = g'(s_0), \quad I(\mathbf{x}+\mathbf{y}) - I(\mathbf{x}) = \langle I'(\mathbf{x} + s_0\mathbf{y}), \mathbf{y}\rangle,$$

where $s_0 \in (0, 1)$ will most likely depend on $\mathbf{x}$ and $\mathbf{y}$. Hence, the quotient Q in (2.29) becomes

$$Q = \langle I'(\mathbf{x} + s_0\mathbf{y}) - I'(\mathbf{x}), \mathbf{y}/\|\mathbf{y}\|\rangle,$$

and

$$0 \le |Q| \le \|I'(\mathbf{x} + s_0\mathbf{y}) - I'(\mathbf{x})\| \to 0$$

as $\|\mathbf{y}\| \to 0$, thanks to the continuity of I'. □

As a consequence of this result, and to avoid confusion with other equivalent definitions, we introduce the concept of a $\mathcal{C}^1$-functional, as another way to identify differentiable functionals. We also define some other usual concepts.

Definition 2.14 Let $I : \mathbb{H} \to \mathbb{R}$ be defined on a Hilbert space $\mathbb{H}$.

1. I is said to be $\mathcal{C}^1$, if it is differentiable (and then the derivative $I' : \mathbb{H} \to \mathbb{H}$ is continuous).
2. I is said to be coercive if

$$\lim_{\|\mathbf{x}\| \to \infty} I(\mathbf{x}) = +\infty.$$

3. I is said to be locally Lipschitz if it is differentiable and

$$\|I'(\mathbf{x}) - I'(\mathbf{y})\| \le M\|\mathbf{x} - \mathbf{y}\|, \quad \mathbf{x}, \mathbf{y} \in \mathbb{K}, M = M(\mathbb{K}) > 0, \tag{2.30}$$

for every bounded subset $\mathbb{K} \subset \mathbb{H}$.

The preceding lemma has, at least, two important consequences that are convenient to highlight just as they are important in finite dimension. The first refers to the orthogonality property of the derivative I' with respect to level sets of I; the second, the relevance of the flow of the derivative I'. The two of them are direct consequences of the chain rule.

Let $I : \mathbb{H} \to \mathbb{R}$ be a differentiable functional, and $\gamma : (-\epsilon, \epsilon) \to \mathbb{H}$ a differentiable curve with continuous tangent vector

$$\gamma'(s) \in \mathbb{H}, \quad s \in (-\epsilon, \epsilon),$$

for some positive ϵ.

Lemma 2.6

1. *The composition*

$$g(t) = I(\gamma(t)) : (-\epsilon, \epsilon) \to \mathbb{R}$$

is differentiable and

$$g'(t) = \langle I'(\gamma(t)), \gamma'(t)\rangle, \quad t \in (-\epsilon, \epsilon).$$

2. *For every* $\mathbf{x} \in \mathbb{H}$*, the vector* $I'(\mathbf{x})$ *is orthogonal to the level set of* I *through* $\mathbf{x}$*.*

3. *Suppose, in addition to the previous assumptions, that* $I' : \mathbb{H} \to \mathbb{H}$ *is coercive and locally Lipschitz Then, for arbitrary initial vector* $\mathbf{x}_0 \in \mathbb{H}$*, the differential infinite-dimensional system*

$$\mathbf{x}'(t) = -I'(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0 \tag{2.31}$$

is defined for every positive time $t > 0$*, and, for every such* $t > 0$*,*

$$\frac{d}{dt} I(\mathbf{x}(t)) = -\|I'(\mathbf{x}(t))\|^2.$$

Moreover, if we write

$$\mathbf{x}'(t; \mathbf{x}_0) = -I'(\mathbf{x}(t; \mathbf{x}_0)), \quad \mathbf{x}(0; \mathbf{x}_0) = \mathbf{x}_0,$$

to stress the dependence of the solution $\mathbf{x}$ *on the initial condition* $\mathbf{x}_0$*, then the mapping* $\mathbf{x}(t; \mathbf{x}_0)$ *is continuous in* $\mathbf{x}_0$*.*

Given the information we already have on differentiable functionals and how everything is similar, almost word by word, with the finite-dimensional setting, the proof of this lemma follows exactly as in that elementary situation. The unique solution of the gradient differential system (2.31) rests, as in the finite-dimensional case, on the contraction principle Theorem 1.1.

Finally, in practice, it may not be possible to clearly see the derivative $I'(\mathbf{x})$ from the computation of the directional derivatives $\mathbf{T}(\mathbf{x}, \mathbf{y})$ given by its definition

$$\mathbf{T}(\mathbf{x}, \mathbf{y}) = \left.\frac{d}{d\epsilon} I(\mathbf{x} + \epsilon \mathbf{y})\right|_{\epsilon=0}.$$

The following basic lemma informs us how to do so.

Lemma 2.7 *For a fixed element* $\mathbf{x} \in \mathbb{H}$*, the unique optimal solution of the quadratic functional*

$$\mathbf{y} \mapsto \frac{1}{2}\|\mathbf{y}\|^2 - \mathbf{T}(\mathbf{x}, \mathbf{y}) \tag{2.32}$$

is precisely $\mathbf{y} = I'(\mathbf{x})$.

Proof This is checked easily by writing

$$\mathbf{T}(\mathbf{x}, \mathbf{y}) = \langle \tilde{\mathbf{T}}\mathbf{x}, \mathbf{y} \rangle,$$

and completing squares

$$\frac{1}{2}\|\mathbf{y}\|^2 - \langle \tilde{\mathbf{T}}\mathbf{x}, \mathbf{y} \rangle = \frac{1}{2}\|\mathbf{y} - \tilde{\mathbf{T}}\mathbf{x}\|^2 - \frac{1}{2}\|\mathbf{x}\|^2.$$

□

2.12 Some Other Important Spaces of Functions

In this chapter, we have introduced the most important spaces of functions utilized in Applied Analysis. There are some important variants that are worth mentioning, and spaces of a different nature. We do not devout more time to study these as they are the subject of more advanced courses in Functional Analysis. There is additional important material in the final Appendix.

1. If (Ω, σ, μ) is an abstract measure space, one can define the corresponding $L^p(d\mu)$-spaces, $1 \leq p \leq \infty$, as one would anticipate

$$L^p(d\mu) = \{f : \Omega \to \mathbb{R}, \sigma - \text{measurable} : \int_\Omega |f(\boldsymbol{x})|^p \, d\mu(\boldsymbol{x}) < \infty\},$$

for p, finite; and the corresponding definition for $p = \infty$. In particular, if $\Omega \subset \mathbb{R}^N$, and $w(\boldsymbol{x}) : \Omega \to \mathbb{R}$ is a positive, integrable function, regarded in this context as a weight, then

$$L^p(w) = \{f : \Omega \to \mathbb{R} : \int_\Omega |f(\boldsymbol{x})|^p w(\boldsymbol{x}) \, d\boldsymbol{x} < \infty\}.$$

Even the spaces ℓ^p may be understood in this context as L^p-spaces with respect to the counting measure in $\mathbb{N}$.

2. So far we have used integration to provide ways to measure the size of functions in suitable classes of measurable functions. These collection of functions are rather huge sets including functions with may exhibit rather irregular behavior. There are ways to restrict attention to regular classes of functions like

$$\mathcal{C}^m(\Omega) = \{f : \Omega \to \mathbb{R} : \text{ with continuous derivatives up to order } m\},$$

with $m \in \mathbb{N}$, and Ω an open subset of $\mathbb{R}^N$. However, given that such functions, or some of their derivatives, can blow up as we approach the boundary of Ω, there does not seem to be a unified way to measure their size only with derivatives and without introducing integration in any way. One possibility is to write

$$p_K(f) = \max_{\boldsymbol{x} \in K, |\alpha| \leq m} |\nabla^\alpha f(\boldsymbol{x})|$$

for a compact set $K \subset \Omega$. This of course does not give a full measure of f as it does not provide information on a function outside the set K. The functions $p_K(f)$ are not norms, but the full collection $p_K(f)$ for an increasing sequence of compact sets $K_j \subset \Omega$ with

$$|\Omega \setminus \cup_j K_j| \to 0, \quad j \to \infty,$$

can be used in such a way that $\mathcal{C}^m(\Omega)$ becomes a topological vector space, something more general than a Banach space. One could try to consider the subspaces

$$\mathcal{C}^m(\overline{\Omega}) = \{f : \overline{\Omega} \to \mathbb{R} : \text{ with continuous derivatives up to } \partial\Omega\},$$

and consider the previous ways of measuring the size of f for $K = \partial\Omega$ (if Ω is bounded).

3. Functions of bounded variation. In a finite interval $J = [x_l, x_r] \subset \mathbb{R}$, we consider the class $\mathcal{P}$ of all finite partitions P of the form

$$x_l = x_0 < x_1 < x_2 < \cdots < x_{k-1} < x_k = x_r,$$

and put

$$V(u) = \sup_{\mathcal{P}} \sum_j |u(x_j) - u(x_{j-1})|.$$

The set of measurable functions u with finite $V(u)$ is the class $BV(J)$ of functions of bounded variation in J. The quantity $V(u) + u(x_l)$ turns out to be a norm in $BV(J)$ under which it becomes a Banach space. The interesting point is that each function $u \in BV(J)$ determines a linear, continuous functional $\mathbf{T}_u$ in $\mathcal{C}(J)$ through the classical Riemann-Stieltjes integral

$$\mathbf{T}_u(v) = \int_J v\,du$$

in such a way that $BV(J)$ becomes the dual of $\mathcal{C}(J)$ under this identification.

4. Complete metric spaces are also a class of objects more general than Banach spaces as they need not be vector spaces where one can take linear combinations. Yet complete metric spaces enjoy some fundamental properties because of the nature of the underlying distance function. As a matter of fact, vector topological spaces that can be shown to be, even locally, metrizable, i.e. their topology comes from the balls of a distance function, do share with metric spaces some of these remarkable properties.
5. Spaces of distributions. We simply mention here the fundamental spaces of distributions because of their crucial role in modern Analysis. There is more information in the final Appendix.

2.13 Exercises

1. Show that $L^1(\Omega)$ and $L^\infty(\Omega)$ are Banach spaces for every open subset $\Omega \subset \mathbb{R}^N$.

2. Prove that if $f_j \to f$ in $L^p(\Omega)$ for $p \geq 1$, and $g \in \mathcal{C}(\overline{\Omega})$ is a continuous function up to the boundary of Ω, then

$$\int_\Omega f_j(\mathbf{x})g(\mathbf{x})\,d\mathbf{x} \to \int_\Omega f(\mathbf{x})g(\mathbf{x})\,d\mathbf{x}.$$

3. Seek a counterexample of a bounded sequence $\{u_j\}$ in $W^{1,1}(0,1)$ which is not equicontinuous.
4. Take

$$u_j(x) = \sin(2^j \pi x), \quad x \in [0,1].$$

   (a) Check that for arbitrary points $0 \leq a < b \leq 1$, we always have

$$\int_a^b u_j(x)\,dx \to 0.$$

   Draw a picture of the graphs of the initial members of the sequence to visualize the cancellation property as $j \to \infty$.

   (b) Calculate the limit of the integrals

$$\int_a^b u_j^2(x)\,dx$$

   as $j \to \infty$ in an arbitrary subinterval $(a,b) \subset [0,1]$. Do you notice something a bit unexpected if these results are interpreted in terms of weak convergence?

5. Let $\{f_j\}$ be a bounded sequence in $L^p(J)$ for $p > 1$, and J a finite interval of $\mathbb{R}$. Suppose there is f in the same space such that

$$\int_a^b f_j(x)\,dx \to \int_a^b f(x)\,dx$$

   for every pair $a < b$ in J. Argue that

$$\int_J g(x)f_j(x)\,dx \to \int_J g(x)f(x)\,dx$$

   for every $g \in L^q(J)$ for the conjugate exponent q.
6. Prove Proposition 2.6.
7. Finish the proof of Proposition 2.7.
8. Prove Proposition 2.8.
9. Prove the second part of Proposition 2.13.

10. For $\mathbb{E}$, the collection of continuous functions in a fixed interval, say, $[0, 1]$, consider the two norms

$$\|f\|_1 = \int_0^1 |f(t)|\,dt, \quad \|f\|_\infty = \sup_{t\in[0,1]} |f(t)|.$$

   (a) Show that $\mathbb{E}_1 = (\mathbb{E}, \|\cdot\|_1)$ is not complete, but $\mathbb{E}_\infty = (\mathbb{E}, \|\cdot\|_\infty)$ is.
   (b) If for each $g \in \mathbb{E}$, we put

$$\langle \mathbf{T}_g, f\rangle = \int_0^1 g(t)f(t)\,dt,$$

   then $\mathbf{T}_g \in \mathbb{E}_1'$. Find its norm.
   (c) Argue that $\mathbf{T}_g \in \mathbb{E}_\infty'$ as well, and

$$\|\mathbf{T}_g\| = \|g\|_1.$$

   (d) Consider the linear functional on $\mathbb{E}$ given by

$$f \mapsto \delta_{1/2}(f) = f(1/2).$$

   Show that $\delta_{1/2} \in \mathbb{E}_\infty'$ but $\delta_{1/2} \notin \mathbb{E}_1'$.
   (e) Prove that

$$\mathbb{H} = \{f \in \mathbb{E} : \int_0^1 f(t)\,dt = 0\}$$

   is a closed hyperplane both in $\mathbb{E}_1'$ and $\mathbb{E}_\infty'$.

11. Let $\mathbb{E}$ be a normed space with norm $\|\cdot\|$.

   (a) Show that the norm is uniformly continuous.
   (b) Every linear map

$$\mathbf{T} : (\mathbb{R}^n, \|\cdot\|_\infty) \to \mathbb{E}$$

   is continuous.
   (c) Every Cauchy sequence is bounded.
   (d) Every norm in $\mathbb{R}^n$ is a continuous function, and there is $M > 0$ with

$$\|\mathbf{x}\| \le M\|\mathbf{x}\|_\infty.$$

   Conclude that every norm in $\mathbb{E}$ is equivalent to the sup-norm, and that all norms in a finite-dimensional vector space are equivalent.

12. (a) If $\mathbf{x} \in \ell^p$ for some $p > 0$, then $\mathbf{x} \in \ell^q$ for all $q > p$, and

$$\lim_{q \to \infty} \|\mathbf{x}\|_q = \|\mathbf{x}\|_\infty.$$

(b) For every bounded and measurable function $f(t)$ for $t \in [a, b]$,

$$\lim_{p \to \infty} \|f\|_p = \|f\|_\infty.$$

13. Let $\mathbb{E}$ be an infinite-dimensional vector space with an algebraic basis $\{\mathbf{e}_i\}$, $i \in I$, in such a way that for each $\mathbf{x} \in \mathbb{E}$, we can write in a unique way

$$\mathbf{x} = \sum_{i \in I} x_i \mathbf{e}_i$$

with only a finite number of the x_i's being non-zero. Define

$$\|\mathbf{x}\|_\infty = \max\{|x_i| : i \in I\}, \quad \|\mathbf{x}\|_1 = \sum_{i \in I} |x_i|.$$

(a) Argue that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are not equivalent.
(b) For an arbitrary norm $\|\cdot\|$ in $\mathbb{E}$, there is always a linear functional which is not continuous.

14. In the vector space

$$\mathbb{K} = \{u \in C^1[0, 1] : u(0) = 0\},$$

we consider

$$\|u\| = \sup_{t \in [0,1]} |u(t) + u'(t)|.$$

Show that it is a norm in $\mathbb{K}$ equivalent to the norm

$$\|u\|_\infty + \|u'\|_\infty.$$

15. Let

$$\mathbf{T} : L^1(\mathbb{R}) \to L^p(\mathbb{R}), \quad 1 < p \le \infty,$$

be linear. Show the equivalence of the following two assertions:

(a) There is some $w \in L^p(\mathbb{R})$ such that

$$\mathbf{T}(u) = u * w, \quad u \in L^1(\mathbb{R});$$

(b) $\mathbf{T}$ is continuous and

$$\mathbf{T}(u) * v = \mathbf{T}(u * v), \quad u, v \in L^1(\mathbb{R}).$$

16. If $\{\mathbf{x}_j\}$ is an orthogonal basis for a separable, Hilbert space $\mathbb{H}$, argue that $\mathbf{x}_j \rightharpoonup \mathbf{0}$ in $\mathbb{H}$.
17. Let the kernel

$$k(\boldsymbol{x}, \boldsymbol{y}) : \Omega \times \Omega \to \mathbb{R},$$

be given by the formula

$$k(\boldsymbol{x}, \boldsymbol{y}) = \sum_i w_i(\boldsymbol{x}) w_i(\boldsymbol{y}), \quad \boldsymbol{x}, \boldsymbol{y} \in \Omega \subset \mathbb{R}^N,$$

where $\{w_i\}$ is a finite collection of smooth functions. Explore if the formula

$$\langle f, g \rangle = \int_{\Omega\times\Omega} f(\boldsymbol{x}) k(\boldsymbol{x}, \boldsymbol{y}) g(\boldsymbol{y})\, d\boldsymbol{x}\, d\boldsymbol{y}$$

defines an inner-product in the space of measurable functions.
18. In the Hilbert space $L^2(-1, 1)$, show that the family of polynomials

$$L_j(t) = \frac{d^j}{dt^j}[(t^2 - 1)^j]$$

is a orthogonal family.[5]

(a) Compute the minimum of the set of numbers

$$\left\{ \int_{-1}^{1} |t^3 - a_2 t^2 - a_1 t - a_0|^2\, dt : a_0, a_1, a_2 \in \mathbb{R} \right\}.$$

(b) Find the maximum of the set of numbers

$$\int_{-1}^{1} t^3 p(t)\, dt$$

where $p(t)$ is subjected to the constraints

$$\int_{-1}^{1} p(t)\, dt = \int_{-1}^{1} t p(t)\, dt = \int_{-1}^{1} t^2 p(t)\, dt = 0, \quad \int_{-1}^{1} p(t)^2\, dt = 1.$$

[5] Except for normalizing constants, these are the Legendre polynomials.

19. Let J be a finite, closed interval of $\mathbb{R}$. A weight $p(t)$ over J is a positive, continuous function such that all monomials t^j are integrable with respect to $p(t)$, i.e.

$$\int_J t^j p(t)\, dt < \infty.$$

Let $P_j(t)$ be the resulting orthogonal family coming from $\{t^j : j = 0, 1, 2, \dots\}$ after applying the Gram-Schmidt process to it with respect to the inner product

$$\langle f, g\rangle = \int_J f(t)g(t)p(t)\, dt.$$

(a) For every j, it is true that

$$\langle tP_{j-1}(t), P_j(t)\rangle = \langle P_j, P_j\rangle.$$

(b) There are two sequences of numbers $\{a_j\}$, $\{b_j\}$, with $b_j > 0$, such that

$$P_j(t) = (t - a_j)P_{j-1}(t) - b_j P_{j-2}(t), \quad j \geq 2.$$

(c) P_j has j different real roots in J.

20. Consider the complex Hilbert space $\mathbb{H}$ of 2π-periodic, complex-valued functions $\mathbf{u} : [-\pi, \pi] \to \mathbb{C}$ under the inner product

$$\langle \mathbf{u}, \mathbf{v}\rangle = \int_{-\pi}^{\pi} \mathbf{u}(t)\overline{\mathbf{v}(t)}\, dt$$

where $\bar{\mathbf{z}}$ stands for the conjugate of the complex number $\mathbf{z}$. Check that the exponential system

$$\{\frac{1}{2\pi}\exp(ijt)\}_{j=0,\pm1,\pm2,\dots}$$

is an orthonormal basis for $\mathbb{H}$.

21. The Haar system. Consider the function

$$\psi(t) = \begin{cases} 1, & 0 \leq t < 1/2, \\ -1, & 1/2 \leq t < 1 \\ 0, & else. \end{cases}$$

For a couple of integers $n, m \in \mathbb{Z}$, define

$$\psi_{m,n}(t) = 2^{n/2}\psi(2^n t - m).$$

(a) Show that this collection of functions is an orthonormal basis of $L^2(\mathbb{R})$.
(b) Show that the family of functions

$$\{1\} \cup \{\psi_{k,n}(t) : n \in \mathbb{N} \cup \{0\}, 0 \leq m < 2^n\}$$

is an orthonormal basis of $L^2(0, 1)$.

22. Wavelets. Let $\psi(t)$ be a real function. For fixed numbers $a > 1$ and $b > 0$, define

$$\psi_{m,n}(t) = \frac{1}{\sqrt{a^m}} \psi\left(\frac{t - nb}{a^m}\right),$$

for $n, m \in \mathbb{N}$. Find conditions on the function ψ in such a way that the family $\{\psi_{m,n}\}$ is an orthonormal basis of $L^2(\mathbb{R})$.
23. Find the orthogonal complement of $H_0^1(J)$ in $H^1(J)$, for an interval J.
24. Use the orthogonality mechanism around (2.27), to find another two families of orthogonal functions in $L^2(0, 1)$.
25. For $J = [0, 1]$, let $\mathbb{E}$ be the set of measurable functions defined on J. Set

$$d(f, g) = \int_0^1 \frac{|f(t) - g(t)|}{1 + |f(t) - g(t)|} \, dt.$$

Show that d is a distance for which $\mathbb{E}$ becomes a complete metric space.
26. Consider the space

$$\mathbb{X} = \{f(x) : (0, 1) \to \mathbb{R}, \text{ measurable} : \int_0^1 |f(x)|^{1/2} \, dx < +\infty\}.$$

(a) Show that $\mathbb{X}$ is a vector space.
(b) Argue that

$$\|f\|_{1/2}^{1/2} = \int_0^1 |f(x)|^{1/2} \, dx$$

is not a norm.
(c) Check that

$$d(f, g) = \int_0^1 |f(x) - g(x)|^{1/2} \, dx$$

is a distance on $\mathbb{X}$ that makes it a complete metric space.

27. (Jordan-Von Neumann Theorem) Let $\mathbb{E}$ be a normed space. Show that the norm comes from a inner product if and only if the parallelogram identity

$$\|\mathbf{x}+\mathbf{y}\|^2+\|\mathbf{x}-\mathbf{y}\|^2=2\|\mathbf{x}\|^2+2\|\mathbf{y}\|^2, \quad \mathbf{x}, \mathbf{y} \in \mathbb{E},$$

holds.
28. Multidimensional Fourier series. Check that the family of functions

$$\{\frac{1}{2\pi}\exp(i\boldsymbol{j}\cdot\boldsymbol{x})\}_{\boldsymbol{j}\in\mathbb{Z}^N}$$

is a orthonormal basis for the Hilbert space $L^2([-\pi,\pi]^N;\mathbb{C})$.
29. A simple obstacle problem. Through the projection theorem Proposition 2.11 to prove that there is a unique solution to the variational problem

$$\text{Minimize in } u \in H_0^1(0,1): \quad \int_0^1 u'(x)^2\,dx$$

subject to $u(x) \le \phi(x)$, where $\phi(x)$ is a continuous function with $\phi(1), \phi(0) < 0$. Write explicitly a condition characterizing the minimizer.
30. For the functional

$$E: H_0^1(0,1) \to \mathbb{R}, \quad E(u)=\int_0^1 [\psi(u'(x))+\phi(u(x))]\,dx$$

where ψ and ϕ are $\mathcal{C}^1$-, real functions such that

$$|\psi(u)|, |\phi(u)| \le C(1+u^2)$$

for a positive constant C, write a quadratic variational problem whose unique solution yields precisely $E'(u)$ for an arbitrary $u \in H_0^1(0,1)$. We will later calculate explicitly such derivative.
31. For $p \in (0,1)$, consider the function

$$d(f,g)=\int_\Omega |f(\boldsymbol{x})-g(\boldsymbol{x})|^p\,d\boldsymbol{x},$$

for $f, g \in \mathcal{C}^\infty(\Omega)$ where $\Omega \subset \mathbb{R}^N$ is an open set as regular as it may be necesssary.

   (a) Check that $(\mathcal{C}^\infty(\Omega), d)$ is a metric space.
   (b) Consider its completion, and show that $L^p(\Omega)$ is a complete metric space under the same distance function.

Chapter 3
Introduction to Convex Analysis: The Hahn-Banach and Lax-Milgram Theorems

3.1 Overview

Before we start diving into integral functionals, it is important to devout some time to understand relevant facts for abstract variational problems. Because in such a situation we do not assume any explicit form of the underlying functional, these results cannot be as fine as those that can be shown when we materialize functionals and spaces. However, the general route to existence of minimizers is essentially the same for all kinds of functionals: it is called the direct method of the Calculus of Variations. This chapter can be considered then as a brief introduction to the fundamental field of Convex Analysis.

The classical Hahn-Banach theorem is one of those basic chapters of Functional Analysis that needs to be known. In particular, it is the basic tool to prove one of the most important existence results of minimizers under convexity and coercivity assumptions in an abstract form. There are several versions of this important theorem: one analytic dealing with the extension of linear functionals; and two geometric that focus on separation principles for convex sets. There are many applications of these fundamental results that are beyond the scope of this text. Some of them will be mentioned in the final Appendix of the book.

We also discuss another two fundamental result that readers ought to know. The first one is basic for quadratic functionals in an abstract, general format: the Lax-Milgram theorem. Its importance in variational problems and Partial Differential Equations cannot be underestimated. The second is also an indispensable tool in Analysis: Stampacchia's theorem for variational inequalities.

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4_3

3.2 The Lax-Milgram Lemma

One first fundamental situation is concerned with quadratic functionals over a general Hilbert space $\mathbb{H}$.

Definition 3.1 Let $\mathbb{H}$ be a (real) Hilbert space.

- A bilinear form on $\mathbb{H}$ is a mapping

$$A(\mathbf{u}, \mathbf{v}) : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$$

  that is linear on each variable separately. If A is symmetric, i.e.

$$A(\mathbf{u}, \mathbf{v}) = A(\mathbf{v}, \mathbf{u}), \quad \mathbf{u}, \mathbf{v} \in \mathbb{H},$$

  the function

$$a(\mathbf{u}) = \frac{1}{2} A(\mathbf{u}, \mathbf{u})$$

  is identified as its associated quadratic form.
- The bilinear form A is continuous if there is a positive constant $C > 0$ with

$$|A(\mathbf{u}, \mathbf{v})| \leq C\|\mathbf{u}\| \, \|\mathbf{v}\|,$$

  for every pair $\mathbf{u}, \mathbf{v} \in \mathbb{H}$. If A is symmetric, then

$$2a(\mathbf{u}) \leq C\|\mathbf{u}\|^2, \quad \mathbf{u} \in \mathbb{H}.$$

- The bilinear form A is coercive if there is a positive constant $c > 0$ such that

$$c\|\mathbf{u}\|^2 \leq A(\mathbf{u}, \mathbf{u})$$

  for every $\mathbf{u} \in \mathbb{H}$. If, in addition, A is symmetric,

$$c\|\mathbf{u}\|^2 \leq 2a(\mathbf{u}), \quad \mathbf{u} \in \mathbb{H}.$$

If $a(\mathbf{u})$ is the quadratic form coming from a certain bilinear, symmetric form, we would like to look at the variational problem

$$\text{Minimize in } \mathbf{u} \in \mathbb{H} : \quad a(\mathbf{u}) - \langle \mathbf{U}, \mathbf{u} \rangle \tag{3.1}$$

for a given $\mathbf{U} \in \mathbb{H}$. The following classic result provides a very clear answer to such quadratic problems.

Theorem 3.1 (Lax-Milgram) *Let $A(\mathbf{u}, \mathbf{v})$ be a symmetric, continuous, coercive bilinear form over a Hilbert space $\mathbb{H}$, with associated quadratic form $a(\mathbf{u})$. The previous variational problem (3.1) admits a unique minimizer $\overline{\mathbf{u}} \in \mathbb{H}$ that is determined by the condition*

$$A(\overline{\mathbf{u}}, \mathbf{v}) = \langle \mathbf{U}, \mathbf{v} \rangle \tag{3.2}$$

for every $\mathbf{v} \in \mathbb{H}$.

Proof We start by making sure that the variational problem (3.1) is well-posed in the sense that if we put

$$m = \inf_{\mathbf{u} \in \mathbb{H}} (a(\mathbf{u}) - \langle \mathbf{U}, \mathbf{u} \rangle)$$

then $m \in \mathbb{R}$. Indeed, by the coercivity property we can write

$$a(\mathbf{u}) - \langle \mathbf{U}, \mathbf{u} \rangle \geq \frac{c}{2} \|\mathbf{u}\|^2 - \langle \mathbf{U}, \mathbf{u} \rangle,$$

and completing squares in the right-hand side,

$$a(\mathbf{u}) - \langle \mathbf{U}, \mathbf{u} \rangle \geq \frac{c}{2} \left\| \mathbf{u} - \frac{1}{c} \mathbf{U} \right\|^2 - \frac{1}{2c} \|\mathbf{U}\|^2.$$

The resulting inequality shows that

$$m \geq -\frac{1}{2c} \|\mathbf{U}\|^2.$$

Let, then, $\{\mathbf{u}^{(j)}\}$ be a minimizing sequence

$$a(\mathbf{u}^{(j)}) - \langle \mathbf{U}, \mathbf{u}^{(j)} \rangle \searrow m.$$

In particular, by the calculations just indicated,

$$\frac{c}{2} \left\| \mathbf{u}^{(j)} - \frac{1}{c} \mathbf{U} \right\|^2 \leq a(\mathbf{u}^{(j)}) - \langle \mathbf{U}, \mathbf{u}^{(j)} \rangle + \frac{1}{2c} \|\mathbf{U}\|^2,$$

which shows that $\{\mathbf{u}^{(j)}\}$ is uniformly bounded in $\mathbb{H}$, and hence, for some subsequence which we do not care to relabel, we will have $\mathbf{u}^{(j)} \rightharpoonup \overline{\mathbf{u}}$ for some $\overline{\mathbf{u}} \in \mathbb{H}$. Again by the coercivity property, we find

$$0 \leq a(\mathbf{u}^{(j)} - \overline{\mathbf{u}}) = \frac{1}{2} A(\mathbf{u}^{(j)}, \mathbf{u}^{(j)}) - A(\mathbf{u}^{(j)}, \overline{\mathbf{u}}) + \frac{1}{2} A(\overline{\mathbf{u}}, \overline{\mathbf{u}}),$$

that is

$$A(\mathbf{u}^{(j)}, \overline{\mathbf{u}}) - \frac{1}{2}A(\overline{\mathbf{u}}, \overline{\mathbf{u}}) \leq a(\mathbf{u}^{(j)}).$$

If we take limits in j, because $A(\cdot, \overline{\mathbf{u}})$ is a linear functional,

$$A(\overline{\mathbf{u}}, \overline{\mathbf{u}}) - \frac{1}{2}A(\overline{\mathbf{u}}, \overline{\mathbf{u}}) \leq \liminf_{j\to\infty} a(\mathbf{u}^{(j)}),$$

and

$$a(\overline{\mathbf{u}}) \leq \liminf_{j\to\infty} a(\mathbf{u}^{(j)}). \tag{3.3}$$

Again, because $\langle \mathbf{U}, \cdot \rangle$ is a linear operation, we can conclude, by the very definition of m as the infimum for our problem, that

$$m \leq a(\overline{\mathbf{u}}) - \langle \mathbf{U}, \overline{\mathbf{u}} \rangle \leq \liminf_{j\to\infty} [a(\mathbf{u}^{(j)}) - \langle \mathbf{U}, \mathbf{u}^{(j)} \rangle] = m.$$

Consequently

$$m = a(\overline{\mathbf{u}}) - \langle \mathbf{U}, \overline{\mathbf{u}} \rangle,$$

and $\overline{\mathbf{u}}$ is truly a minimizer.

For the second part of the proof, we proceed in two steps. We first show that (3.2) holds indeed for the previous minimizer $\overline{\mathbf{u}}$. To this aim, take $\mathbf{v} \in \mathbb{H}$, and consider the one-dimensional section

$$g(\epsilon) = I(\overline{\mathbf{u}} + \epsilon\mathbf{v}), \quad I(\mathbf{u}) = a(\mathbf{u}) - \langle \mathbf{U}, \mathbf{v} \rangle.$$

Because $\overline{\mathbf{u}}$ is a (global) minimizer for I, $\epsilon = 0$ ought to be a global minimizer for g. On the other hand, it turns out, taking advantage of the bilinearity and symmetry of A, that g is quadratic, and so smooth. Consequently, $g'(0) = 0$, and a quick computation yields

$$0 = g'(0) = A(\overline{\mathbf{u}}, \mathbf{v}) - \langle \mathbf{U}, \mathbf{v} \rangle.$$

The arbitrariness of $\mathbf{v}$ leads to (3.2).

Secondly, we argue that there cannot be two different vectors $\mathbf{u}_i$, $i = 1, 2$, for which (3.2) holds. If it were true that

$$A(\mathbf{u}_i, \mathbf{v}) = \langle \mathbf{U}, \mathbf{v} \rangle, \quad \mathbf{v} \in \mathbb{H}, i = 1, 2,$$

we would definitely have

$$A(\mathbf{u}_1 - \mathbf{u}_2, \mathbf{v}) = 0$$

for every $\mathbf{v} \in \mathbb{H}$; in particular, for $\mathbf{v} = \mathbf{u}_1 - \mathbf{u}_2$, we would find

$$a(\mathbf{u}_1 - \mathbf{u}_2) = 0.$$

The coercivity of the quadratic form a means that $\mathbf{u}_1 = \mathbf{u}_2$. □

The Lax-Milgram lemma is a very powerful tool to deal with quadratic functionals and linear optimality conditions (linear differential equations) because it does not prejudge the nature of the underlying Hilbert space $\mathbb{H}$: it is valid for every Hilbert space, and every coercive quadratic functional. The Lax-Milgram lemma can also be shown through Stampachia's theorem (Exercise 10) that is treated in the final section of the chapter.

In the previous proof, we have already anticipated some of the key ideas that will guide us for more general situations.

Example 3.1 The first example is mandatory: take

$$A(\mathbf{u}, \mathbf{v}) = \langle \mathbf{u}, \mathbf{v} \rangle$$

the inner product itself. In this case the variational problem (3.1) becomes the one examined in (2.32). There is nothing to be added.

Example 3.2 The prototypical example, in dimension one, that can be treated under the Lax-Milgram lemma is the following. Let $\mathbb{H}$ be the Hilbert subspace $H_0^1(J)$ of functions in $H^1(J)$ vanishing at the two end-points of J, for a finite interval $J = (x_0, x_1)$ of $\mathbb{R}$. Consider the bilinear form

$$A(u, v) = \int_J [\alpha(x)u'(x)v'(x) + \beta(x)u(x)v(x)]\,dx,$$

where functions $\alpha(x)$ and $\beta(x)$ will be further restricted in the sequel. A is definitely symmetric. It is bounded if both functions α, and β belong to $L^\infty(J)$. It is coercive if the same two functions are positive and bounded away from zero. The application of Theorem 3.1 directly leads to the following result.

Corollary 3.1 *Suppose there is a constant $C > 0$ such that*

$$0 < C < \min_{x \in J}(\alpha(x), \beta(x)) \le \max_{x \in J}(\alpha(x), \beta(x)) \le \frac{1}{C}.$$

For every function $\gamma(x) \in L^2(J)$, the variational problem

$$\textit{Minimize in } u(x) \in H_0^1(J): \quad \int_J [\frac{1}{2}\alpha(x)u'(x)^2 + \frac{1}{2}\beta(x)u(x)^2 - \gamma(x)u(x)]\,dx$$

admits a unique minimizer $\overline{u}$ which is characterized by the condition

$$\int_J [\alpha(x)\overline{u}'(x)v'(x) + \beta(x)\overline{u}(x)v(x) - \gamma(x)v(x)]\,dx = 0 \tag{3.4}$$

for every $v \in H_0^1(J)$.

This last integral condition is typically called the weak form of the boundary-value (Sturm-Liouville) problem

$$-[\alpha(x)u'(x)]' + \beta(x)u(x) = \gamma(x) \text{ in } J, \quad u(x_0) = u(x_1) = 0. \tag{3.5}$$

3.3 The Hahn-Banach Theorem: Analytic Form

The analytic form of the Hahn-Banach theorem is typically based on Zorn's lemma, which is taken for granted here. The relevant concepts to understand Zorn's lemma are gathered in the following statement.

Definition 3.2 Let **P** be a partially ordered set.

- A subset **Q** is totally ordered if for every pair **u**, **v** in **Q**, either $\mathbf{u} \leq \mathbf{v}$ or $\mathbf{v} \leq \mathbf{u}$.
- An element $\mathbf{w} \in \mathbf{P}$ is an upper bound for a subset $\mathbf{Q} \subset \mathbf{P}$ if $\mathbf{u} \leq \mathbf{w}$ for every $\mathbf{u} \in \mathbf{Q}$.
- An element $\mathbf{m} \in \mathbf{P}$ is maximal if there is no element $\mathbf{u} \in \mathbf{P}$ with $\mathbf{m} \leq \mathbf{u}$ except **m** itself.
- **P** is said to be inductive if every chain (i.e. totally ordered subset) **Q** in **P** has an upper bound.

With these concepts we have:

Lemma 3.1 (Zorn's Lemma) *Every non-empty, inductive, ordered set has a maximal element.*

In addition, we need to define what is understood by a seminorm in a (real) vector space $\mathbb{E}$. It is a real, non-negative function complying with N2 and N3, but not with N1, in Definition 2.1 in Chap. 2.

Definition 3.3 A function

$$p(\mathbf{x}) : \mathbb{E} \to \mathbb{R}$$

is called a subnorm in $\mathbb{E}$ if:

N2 for every pair $\mathbf{x}, \mathbf{y} \in \mathbb{E}$,

$$p(\mathbf{x} + \mathbf{y}) \le p(\mathbf{x}) + p(\mathbf{y});$$

N3' for every positive scalar λ and vector $\mathbf{x} \in \mathbb{E}$,

$$p(\lambda \mathbf{x}) = \lambda p(\mathbf{x}).$$

If, in addition, p is non-negative, then p is called a seminorm.

We are now ready to prove the analytic form of the Hahn-Banach theorem.

Theorem 3.2 *Let p be a subnorm over a (real) vector space $\mathbb{E}$; $\mathbb{M}$, a subspace of $\mathbb{E}$, and T_0 a linear functional on $\mathbb{M}$ with*

$$T_0(\mathbf{x}) \le p(\mathbf{x}), \quad \mathbf{x} \in \mathbb{M}.$$

Then a linear extension T of T_0 to all of $\mathbb{E}$ can be found, i.e. $T|_{\mathbb{M}} = T_0$, so that

$$T(\mathbf{x}) \le p(\mathbf{x}), \quad \mathbf{x} \in \mathbb{E}. \tag{3.6}$$

Proof The standard proof of this result proceeds in two steps. The first one deals with the case in which

$$\mathbb{E} = \mathbb{M} \oplus \langle \mathbf{x}_0 \rangle,$$

where $\langle \mathbf{x}_0 \rangle$ stands for the subspace spanned by some $\mathbf{x}_0 \in \mathbb{E}$. In this situation, every $\mathbf{x} \in \mathbb{E}$ can be written in a unique way in the form $\mathbf{m} + \lambda \mathbf{x}_0$, and, by linearity,

$$T(\mathbf{x}) = T(\mathbf{m}) + \lambda T(\mathbf{x}_0) = T_0(\mathbf{m}) + \lambda T(\mathbf{x}_0).$$

Thus, it suffices to specify how to select the number $\mu = T(\mathbf{x}_0)$ appropriately to ensure inequality (3.6) in this case, that is to say

$$T_0(\mathbf{m}) + \lambda \mu \le p(\mathbf{m} + \lambda \mathbf{x}_0),$$

or

$$\lambda \mu \le p(\mathbf{m} + \lambda \mathbf{x}_0) - T_0(\mathbf{m})$$

for every $\mathbf{m} \in \mathbb{M}$ and every $\lambda \in \mathbb{R}$. For $\lambda > 0$, this last inequality becomes

$$\mu \le p(\mathbf{x}_0 + (1/\lambda)\mathbf{m}) - T_0((1/\lambda)\mathbf{m});$$

while for $\lambda < 0$, we should have

$$\mu \geq -p(-\mathbf{x}_0 - (1/\lambda)\mathbf{m}) + T_0(-(1/\lambda)\mathbf{m}).$$

If we let

$$\mathbf{v} = -(1/\lambda)\mathbf{m}, \quad \mathbf{u} = (1/\lambda)\mathbf{m},$$

it suffices to guarantee that there is a real number μ such that

$$T_0(\mathbf{v}) - p(\mathbf{v} - \mathbf{x}_0) \leq \mu \leq -T_0(\mathbf{u}) + p(\mathbf{u} + \mathbf{x}_0)$$

for every $\mathbf{u}, \mathbf{v} \in \mathbb{M}$. But for every such arbitrary pair, by hypothesis and the subadditivity property of p,

$$T_0(\mathbf{u}) + T_0(\mathbf{v}) = T_0(\mathbf{u} + \mathbf{v}) \leq p(\mathbf{u} + \mathbf{v}) \leq p(\mathbf{u} + \mathbf{x}_0) + p(\mathbf{v} - \mathbf{x}_0).$$

This finishes our first step.

For the second step, consider the class

$$\begin{aligned} \mathcal{G} = \{(\mathbb{F}, T_{\mathbb{F}}) : \mathbb{F} \text{ is subspace}, \mathbb{M} \subset \mathbb{F} \subset \mathbb{E}, T_{\mathbb{F}} \text{ is linear in } \mathbb{F}, \\ T_{\mathbb{F}}|_{\mathbb{M}} = T_0, T_{\mathbb{F}}(\mathbf{x}) \leq p(\mathbf{x}), \mathbf{x} \in \mathbb{F}\}. \end{aligned}$$

$\mathcal{G}$ is an ordered set under the order relation

$$(\mathbb{F}_1, T_{\mathbb{F}_1}) < (\mathbb{F}_2, T_{\mathbb{F}_2}) \text{ when } \mathbb{F}_1 \subset \mathbb{F}_2, \ T_{\mathbb{F}_2}\big|_{\mathbb{F}_1} = T_{\mathbb{F}_1}.$$

$\mathcal{G}$ is non-empty because $(\mathbb{M}, T_0) \in \mathcal{G}$. It is inductive too. To this end, let $(\mathbb{F}_i, T_{\mathbb{F}_i})$ be a chain in $\mathcal{G}$. Set $\mathbb{F} = \cup_i \mathbb{F}_i$ which is a subspace of $\mathbb{E}$ as well. For $\mathbf{x} \in \mathbb{F}$, define

$$T_{\mathbb{F}}(\mathbf{x}) = T_{\mathbb{F}_i}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{F}_i.$$

There is no ambiguity in this definition, in case $\mathbf{x}$ belongs to several of the $\mathbb{F}_i$'s. By Zorn's lemma, there is a maximal element $(\mathbb{H}, T_{\mathbb{H}})$ in $\mathcal{G}$. Suppose $\mathbb{H}$ is not all of $\mathbb{E}$, and let $\mathbf{x}_0 \in \mathbb{E} \setminus \mathbb{H}$. For our first step applied to the direct sum

$$\mathbb{G} = \mathbb{H} \oplus \langle \mathbf{x}_0 \rangle,$$

we could find $T_{\mathbb{G}}$ with

$$T_{\mathbb{G}}|_{\mathbb{H}} = T_{\mathbb{H}}, \quad T_{\mathbb{G}}(\mathbf{x}) \leq p(\mathbf{x}), \mathbf{x} \in \mathbb{G}.$$

This would contradict the maximality of $(\mathbb{H}, T_{\mathbb{H}})$ in $\mathcal{G}$, and, hence $\mathbb{H}$ must be the full space $\mathbb{E}$. $\square$

Among the most important consequences of Theorem 3.2 are the following.

Corollary 3.2 *Let $\mathbb{E}$ be a Banach space, with dual $\mathbb{E}'$.*

1. *Let $\mathbb{F}$ be a linear subspace of $\mathbb{E}$. If $T_0 : \mathbb{F} \to \mathbb{R}$ is linear and continuous ($T_0 \in \mathbb{F}'$), there is $T \in \mathbb{E}'$ with*

$$T|_{\mathbb{F}} = T_0, \quad \|T\| = \|T_0\|.$$

2. *For every non-vanishing vector $\mathbf{x}_0 \in \mathbb{E}$, and real α, there is $T_0 \in \mathbb{E}'$ with*

$$\langle T_0, \mathbf{x}_0 \rangle = \alpha, \quad |\langle T_0, \mathbf{x} \rangle| \le \frac{|\alpha|}{\|\mathbf{x}_0\|} \|\mathbf{x}\|, \quad \mathbf{x} \in \mathbb{E}.$$

In particular, for $\alpha = \|\mathbf{x}_0\|^2$ there is $T_0 \in \mathbb{E}'$ with

$$\langle T_0, \mathbf{x}_0 \rangle = \|\mathbf{x}_0\|^2, \quad \|T_0\| = \|\mathbf{x}_0\|.$$

3. *For every $\mathbf{x} \in \mathbb{E}$,*

$$\|\mathbf{x}\| = \sup_{T \in \mathbb{E}' : \|T\|=1} \langle T, \mathbf{x} \rangle = \max_{T \in \mathbb{E}' : \|T\|=1} \langle T, \mathbf{x} \rangle.$$

Proof For the first part, apply directly Theorem 3.2 to the choice

$$p(\mathbf{x}) = \|T_0\| \, \|\mathbf{x}\|.$$

Take, for the second, $\mathbb{M} = \langle \mathbf{x}_0 \rangle$ and

$$T_0(\lambda \mathbf{x}_0) = \lambda \alpha.$$

Then, it is obvious that

$$T_0(\lambda \mathbf{x}_0) \le \frac{|\alpha|}{\|\mathbf{x}_0\|} \|\lambda \mathbf{x}_0\|.$$

Apply Theorem 3.2 with

$$p(\mathbf{x}) = \frac{|\alpha|}{\|\mathbf{x}_0\|} \|\mathbf{x}\|.$$

The third statement is a consequence of the second one, and is left as an exercise.

□

3.4 The Hahn-Banach Theorem: Geometric Form

A main tool for the study of the geometric form of the Hahn-Banach theorem and its consequences concerning the separation of convex sets is the Minkowski functional of an open convex set. Recall the following.

Definition 3.4

1. A convex set $\mathbf{C}$ of a vector space $\mathbb{E}$ is such that convex combinations of elements of $\mathbf{C}$ stay in $\mathbf{C}$

$$t\mathbf{x} + (1-t)\mathbf{y} \in \mathbf{C}, \quad \mathbf{x}, \mathbf{y} \in \mathbf{C}, t \in [0, 1].$$

2. The convexification of a set $\mathbf{C}$ is given by

$$\text{co}(\mathbf{C}) = \{ \sum_{i, finite} \lambda_i \mathbf{x}_i : \mathbf{x}_i \in \mathbf{C}, \lambda_i \in [0, 1], \sum_i \lambda_i = 1\}.$$

It is immediate to check that the convexification of any set is always convex.

Definition 3.5 For a subset $\mathbf{C} \subset \mathbb{E}$, we define the Minkowski functional of $\mathbf{C}$ as

$$p_{\mathbf{C}}(\mathbf{x}) : \mathbb{E} \to [0, +\infty], \quad p_{\mathbf{C}}(\mathbf{x}) = \inf\{\rho > 0 : \mathbf{x} \in \rho\mathbf{C}\}.$$

Important properties of $p_{\mathbf{C}}$ depend on conditions on $\mathbf{C}$.

Lemma 3.2 *Let $\mathbf{C} \subset \mathbb{E}$ be open, convex, and $\mathbf{0} \in \mathbf{C}$. Then $p_{\mathbf{C}}$ is a subnorm (Definition 3.3), there is a positive constant M with*

$$0 \leq p_{\mathbf{C}}(\mathbf{x}) \leq M\|\mathbf{x}\|, \quad \mathbf{x} \in \mathbb{E},$$

and

$$\mathbf{C} = \{\mathbf{x} \in \mathbb{E} : p_{\mathbf{C}}(\mathbf{x}) < 1\}.$$

Proof Property N3' in Definition 3.3 is immediate. If $\mathbf{C}$ is open, and $\mathbf{0} \in \mathbf{C}$, there is some positive radius $r > 0$ such that the closed ball $\mathbf{B}_r \subset \mathbf{C}$. This means that

$$\mathbf{x} \in \frac{\|\mathbf{x}\|}{r}\mathbf{B}_r \subset \frac{\|\mathbf{x}\|}{r}\mathbf{C},$$

and, by definition,

$$0 \leq p_{\mathbf{C}}(\mathbf{x}) \leq M\|\mathbf{x}\|, \quad M = 1/r.$$

Suppose $\mathbf{x} \in \mathbf{C}$. Because $\mathbf{C}$ is open, for some $\epsilon > 0$, $(1+\epsilon)\mathbf{x} \in \mathbf{C}$, i.e.

$$p_{\mathbf{C}}(\mathbf{x}) \leq \frac{1}{1+\epsilon} < 1.$$

If, on the other hand, $p_{\mathbf{C}}(\mathbf{x}) < 1$, $\rho \in (0, 1)$ can be found so that

$$\mathbf{x} \in \rho\mathbf{C}, \quad \frac{1}{\rho}\mathbf{x} \in \mathbf{C}.$$

Since

$$\mathbf{0}, (1/\rho)\mathbf{x} \in \mathbf{C},$$

by convexity

$$\mathbf{x} = \rho(1/\rho)\mathbf{x} + (1-\rho)\mathbf{0} \in \mathbf{C}.$$

We finally check that $p_{\mathbf{C}}$ complies with N2. We already know that for positive ϵ and $\mathbf{x} \in \mathbb{E}$,

$$\frac{1}{p_{\mathbf{C}}(\mathbf{x}) + \epsilon}\mathbf{x} \in \mathbf{C},$$

because

$$p_{\mathbf{C}}\left(\frac{1}{p_{\mathbf{C}}(\mathbf{x}) + \epsilon}\mathbf{x}\right) < 1.$$

For every couple $\mathbf{x}, \mathbf{y} \in \mathbb{E}$, and $\epsilon > 0$, by the convexity of $\mathbf{C}$, we will have that

$$\frac{t}{p_{\mathbf{C}}(\mathbf{x}) + \epsilon}\mathbf{x} + \frac{1-t}{p_{\mathbf{C}}(\mathbf{y}) + \epsilon}\mathbf{y} \in \mathbf{C},$$

for every $t \in [0, 1]$, i.e.

$$p_{\mathbf{C}}\left(\frac{t}{p_{\mathbf{C}}(\mathbf{x}) + \epsilon}\mathbf{x} + \frac{1-t}{p_{\mathbf{C}}(\mathbf{y}) + \epsilon}\mathbf{y}\right) < 1.$$

By selecting t in such a way that

$$\frac{t}{p_{\mathbf{C}}(\mathbf{x}) + \epsilon} = \frac{1-t}{p_{\mathbf{C}}(\mathbf{y}) + \epsilon}, \quad t = \frac{p_{\mathbf{C}}(\mathbf{x}) + \epsilon}{p_{\mathbf{C}}(\mathbf{x}) + p_{\mathbf{C}}(\mathbf{y}) + 2\epsilon},$$

we will have

$$p_{\mathbf{C}}(\mathbf{x}+\mathbf{y}) \le p_{\mathbf{C}}(\mathbf{x}) + p_{\mathbf{C}}(\mathbf{y}) + 2\epsilon.$$

The arbitrariness of ϵ implies the triangular inequality. □

We now introduce the concept of separation of sets.

Definition 3.6 Let $\mathbb{E}$ be a Banach space with dual $\mathbb{E}'$, and $\mathbf{F}$, $\mathbf{G}$, two subsets of $\mathbb{E}$. For $T \in \mathbb{E}'$, the hyperplane

$$\mathbb{M} = \{T = \alpha\}, \quad \alpha \in \mathbb{R},$$

separates $\mathbf{F}$ from $\mathbf{G}$ if

$$\langle T, \mathbf{x}\rangle \le \alpha, \mathbf{x} \in \mathbf{F}, \quad \langle T, \mathbf{x}\rangle \ge \alpha, \mathbf{x} \in \mathbf{G}.$$

The separation is strict if $\epsilon > 0$ can be found such that

$$\langle T, \mathbf{x}\rangle \le \alpha - \epsilon, \mathbf{x} \in \mathbf{F}, \quad \langle T, \mathbf{x}\rangle \ge \alpha + \epsilon, \mathbf{x} \in \mathbf{G}.$$

One preliminary step corresponds to the particular case in which one of the sets in a singleton.

Lemma 3.3 *Let* $\mathbf{C}$ *be a non-empty, open, convex set in* $\mathbb{E}$*, and* $\mathbf{x}_0 \notin \mathbf{C}$. *Then a functional* $T \in \mathbb{E}'$ *can be found with*

$$\langle T, \mathbf{x}\rangle < \langle T, \mathbf{x}_0\rangle, \quad \mathbf{x} \in \mathbf{C}.$$

In this way, the closed hyperplane $\{T = \langle T, \mathbf{x}_0\rangle\}$ *separates* $\mathbf{C}$ *from* $\{\mathbf{x}_0\}$.

Proof Without loss of generality through a translation, we may suppose that $\mathbf{0} \in \mathbf{C}$. Let $p = p_{\mathbf{C}}$ be the Minkowski functional for $\mathbf{C}$, and let $\mathbb{G}$ be the one-dimensional subspace generated by $\mathbf{x}_0$. Define

$$T_0 \in \mathbb{G}', \quad \langle T_0, \mathbf{x}_0\rangle = 1.$$

Then, for $t > 0$, we have

$$\langle T_0, t\mathbf{x}_0\rangle = t \le tp(\mathbf{x}_0) = p(t\mathbf{x}_0)$$

because $\mathbf{x}_0 \notin \mathbf{C}$; whereas for $t \le 0$,

$$\langle T_0, t\mathbf{x}_0\rangle = t \le 0 \le p(t\mathbf{x}_0).$$

The analytic form of the Hahn-Banach theorem implies that T_0 can be extended to all of $\mathbb{E}$ in such a way that

$$\langle T, \mathbf{x} \rangle \leq p(\mathbf{x}), \quad \mathbf{x} \in \mathbb{E},$$

if $T \in \mathbb{E}'$ is such an extension. Note that T is continuous because, for some $M > 0$,

$$p(\mathbf{x}) \leq M \|\mathbf{x}\|, \quad \mathbf{x} \in \mathbb{E}.$$

In particular, for every $\mathbf{x} \in \mathbf{C}$,

$$\langle T, \mathbf{x} \rangle \leq p(\mathbf{x}) < 1 = \langle T, \mathbf{x}_0 \rangle.$$

□

We are now ready to deal with the first form of the geometric version of the Hahn-Banach theorem.

Theorem 3.3 *Let* $\mathbf{F}$, $\mathbf{G}$ *be two convex, disjoint subsets of the Banach space* $\mathbb{E}$ *with dual* $\mathbb{E}'$. *If at least one of the two is open, there is a closed hyperplane separating them from each other.*

Proof Suppose $\mathbf{F}$ is open, and put

$$\mathbf{C} = \mathbf{F} - \mathbf{G} = \{\mathbf{x} - \mathbf{y} : \mathbf{x} \in \mathbf{F}, \mathbf{y} \in \mathbf{G}\} = \cup_{\mathbf{y} \in \mathbf{G}} (\mathbf{F} - \mathbf{y}).$$

It is elementary to check that $\mathbf{C}$ is convex, open (because each $\mathbf{F} - \mathbf{y}$ is), and $\mathbf{0} \notin \mathbf{C}$. By our Lemma 3.3, conclude that there is $T \in \mathbb{E}'$ such that

$$\langle T, \mathbf{z} \rangle < 0, \quad \mathbf{z} \in \mathbf{C}.$$

Since every difference

$$\mathbf{z} = \mathbf{x} - \mathbf{y}, \quad \mathbf{x} \in \mathbf{F}, \mathbf{y} \in \mathbf{G},$$

belong to $\mathbf{C}$, we find

$$\langle T, \mathbf{x} \rangle - \langle T, \mathbf{y} \rangle < 0.$$

If $\rho \in \mathbb{R}$ is chosen in such a way that

$$\sup_{\mathbf{x} \in \mathbf{F}} \langle T, \mathbf{x} \rangle \leq \rho \leq \inf_{\mathbf{y} \in \mathbf{G}} \langle T, \mathbf{y} \rangle,$$

we will have that the hyperplane $\{T = \rho\}$ separates $\mathbf{F}$ from $\mathbf{G}$. □

Asking for a strict separation between $\mathbf{F}$ and $\mathbf{G}$ demands much more restrictive assumptions on both sets.

Theorem 3.4 *Let* $\mathbf{F}$, $\mathbf{G}$ *be two disjoint, convex sets of* $\mathbb{E}$. *Suppose* $\mathbf{F}$ *is closed, and* $\mathbf{G}$*, compact. There is, then, a closed hyperplane that strictly separates them from each other.*

Proof We also consider the difference set $\mathbf{C}$ as in the proof of the first version. We leave as an exercise to show that $\mathbf{C}$ is, in addition to convex, closed. The more restrictive hypotheses are to be used in a fundamental way here. Since $\mathbf{0} \notin \mathbf{C}$, there is some positive r such that the open ball $\mathbf{B}_r$, centered at $\mathbf{0}$ y radius r, has an empty intersection with $\mathbf{C}$. By Theorem 3.3 applied to $\mathbf{C}$ and $\mathbf{B}_r$, we find a non-null $T \in \mathbb{E}'$ so that

$$\langle T, \mathbf{x} - \mathbf{y} \rangle \le \langle T, r\mathbf{z} \rangle, \quad \mathbf{x} \in \mathbf{F}, \mathbf{y} \in \mathbf{G}, \mathbf{z} \in \mathbf{B}_1.$$

This yields, by taking the infimum in $\mathbf{z} \in \mathbf{B}_1$,

$$\frac{1}{r} \langle T, \mathbf{x} - \mathbf{y} \rangle \le -\|T\|, \quad \mathbf{x} \in \mathbf{F}, \mathbf{y} \in \mathbf{G},$$

that is to say

$$\langle T, \mathbf{x} \rangle + \frac{r\|T\|}{2} \le \langle T, \mathbf{y} \rangle - \frac{r\|T\|}{2}, \quad \mathbf{x} \in \mathbf{F}, \mathbf{y} \in \mathbf{G}.$$

As before, for ρ selected to ensure that

$$\sup_{\mathbf{x} \in \mathbf{F}} \langle T, \mathbf{x} \rangle + \frac{r\|T\|}{2} \le \rho \le \inf_{\mathbf{y} \in \mathbf{G}} \langle T, \mathbf{y} \rangle - \frac{r\|T\|}{2},$$

we will have that the hyperplane $\{T = \rho\}$ strictly separates $\mathbf{F}$ from $\mathbf{G}$. □

3.5 Some Applications

Our first application is a helpful fact that sometimes may have surprising consequences. Its proof is a direct application of Theorem 3.4 when the closed set is a closed subspace, and the compact set is a singleton.

Corollary 3.3 *Let* $\mathbb{M}$ *be a subspace of a Banach space* $\mathbb{E}$ *with dual* $\mathbb{E}'$. *If the closure* $\overline{\mathbb{M}}$ *of* $\mathbb{M}$ *is not the full space* $\mathbb{E}$*, then there is a non-null* $T \in \mathbb{E}'$ *such that*

$$\langle T, \mathbf{x} \rangle = 0, \quad \mathbf{x} \in \mathbb{M}.$$

Proof For the choice

$$\mathbf{F} = \overline{\mathbb{M}}, \quad \mathbf{G} = \{\mathbf{x}_0\}, \mathbf{x}_0 \in \mathbb{E} \setminus \overline{\mathbb{M}},$$

by Theorem 3.4, we would find $T \in \mathbb{E}'$ and $\rho \in \mathbb{R}$ such that

$$\langle T, \mathbf{x} \rangle < \rho < \langle T, \mathbf{x}_0 \rangle, \quad \mathbf{x} \in \mathbb{M}.$$

Since

$$\lambda \mathbf{x} \in \mathbb{M} \text{ if } \mathbf{x} \in \mathbb{M},$$

we realize that the only value of ρ compatible with the previous left-hand side inequality is $\rho = 0$, and then the inequality must be an equality. □

One of the most appealing applications of this corollary consists in the conclusion that a subspace $\mathbb{M}$ is dense in $\mathbb{E}$, $\overline{\mathbb{M}} = \mathbb{E}$, if the assumption

$$\langle T, \mathbf{x} \rangle = 0 \text{ for every } \mathbf{x} \in \mathbb{M}$$

implies that $T \equiv \mathbf{0}$ as elements of $\mathbb{E}'$.

There is a version of the preceding corollary for general convex sets, not necessarily subspaces, that is usually utilized in a negative form.

Corollary 3.4 *Let* $\mathbf{F}$ *be a convex set of a Banach space* $\mathbb{E}$ *with dual* $\mathbb{E}'$. *If a further set* $\mathbf{G}$ *cannot be separated from* $\mathbf{F}$ *in the sense*

$$\langle T, \mathbf{x} \rangle + \rho \geq 0 \textit{ for every } \mathbf{x} \in \mathbf{F} \textit{ and some } T \in \mathbb{E}',$$

implies

$$\langle T, \mathbf{x} \rangle + \rho \geq 0 \textit{ for every } \mathbf{x} \in \mathbf{G},$$

then $\mathbf{G} \subset \overline{\mathbf{F}}$.

There is even a version of this result in the dual of a Banach space.

Corollary 3.5 *Let* $\mathbf{F}$ *be a convex set of the dual space* $\mathbb{E}'$ *of a Banach space* $\mathbb{E}$. *If a further subset* $\mathbf{G} \subset \mathbb{E}'$ *cannot be separated from* $\mathbf{F}$ *in the sense*

$$\langle T, \mathbf{x} \rangle + \rho \geq 0 \textit{ for every } T \in \mathbf{F} \textit{ and some } \mathbf{x} \in \mathbb{E}$$

implies

$$\langle T, \mathbf{x} \rangle + \rho \geq 0 \textit{ for every } T \in \mathbf{G},$$

then $\mathbf{G} \subset \overline{\mathbf{F}}$.

The following form is, however, usually better adapted to practical purposes.

Corollary 3.6 *Let* $\mathbf{F}$ *be a set of the dual space* $\mathbb{E}'$ *of a Banach space* $\mathbb{E}$*. If a further subset* $\mathbf{G} \subset \mathbb{E}'$ *cannot be separated from* $\mathbf{F}$ *in the following sense: whenever*

$$\langle T, \mathbf{x}\rangle + \rho < 0 \textit{ for some } T \in \mathbf{G}, \mathbf{x} \in \mathbb{E}, \rho \in \mathbb{R},$$

then there is $\hat{T} \in \mathbf{F}$ *with*

$$\langle \hat{T}, \mathbf{x}\rangle + \rho < 0;$$

then $\mathbf{G} \subset \overline{\text{co}(\mathbf{F})}$.

A fundamental consequence for variational methods is the following.

Theorem 3.5 *Every convex, closed set* $\mathbf{C}$ *in a Banach space* $\mathbb{E}$ *is weakly closed.*

Proof Let

$$\mathbf{u}_j \rightharpoonup \mathbf{u}, \quad \mathbf{u}_j \in \mathbf{C},$$

with $\mathbf{C}$ convex and closed in $\mathbb{E}$. Suppose $\mathbf{u} \notin \mathbf{C}$. We can apply Theorem 3.4, and conclude the existence of an element $T \in \mathbb{E}'$ and some $\epsilon > 0$ such that

$$\langle T, \mathbf{u}\rangle < \epsilon, \quad \langle T, \mathbf{v}\rangle \ge \epsilon,$$

for all $\mathbf{v} \in \mathbf{C}$. This is a contradiction because each $\mathbf{v} = \mathbf{u}_j \in \mathbf{C}$, and then we would have

$$\langle T, \mathbf{u}_j\rangle \ge \epsilon, \quad \langle T, \mathbf{u}\rangle < \epsilon,$$

which is impossible if

$$\lim_{j\to\infty} \langle T, \mathbf{u}_j\rangle = \langle T, \mathbf{u}\rangle.$$

$\square$

3.6 Convex Functionals, and the Direct Method

The Lax-Milgram lemma Theorem 3.1 is a very efficient tool to deal with quadratic functionals and linear boundary-value problems for PDEs. If more general situations are to be examined, one needs to rely in more general assumptions, and convexity stands as a major structural property. We gather here the main concepts related to minimization of abstract functionals.

Definition 3.7 A functional $I : \mathbb{E} \to \mathbb{R}$, defined in a Banach space $\mathbb{E}$, is:

1. convex if

$$I(t\mathbf{u}_1 + (1-t)\mathbf{u}_0) \le I(\mathbf{u}_1) + (1-t)I(\mathbf{u}_0)$$

for every

$$\mathbf{u}_1, \mathbf{u}_0 \in \mathbb{E}, \quad t \in [0,1];$$

it is strictly convex if, in addition, the identity

$$I(t\mathbf{u}_1 + (1-t)\mathbf{u}_0) = I(\mathbf{u}_1) + (1-t)I(\mathbf{u}_0)$$

is only possible when

$$\mathbf{u}_1 = \mathbf{u}_0, \quad t(1-t) = 0;$$

2. lower semicontinuous if

$$I(\mathbf{u}) \le \liminf_{j\to\infty} I(\mathbf{u}_j)$$

whenever $\mathbf{u}_j \to \mathbf{u}$ in $\mathbb{E}$;
3. coercive if

$$I(\mathbf{u}) \to \infty \text{ when } \|\mathbf{u}\| \to \infty;$$

4. bounded from below if there is a constant $c \in \mathbb{R}$ such that

$$c \le I(\mathbf{u}), \quad \mathbf{u} \in \mathbb{E}.$$

It is not difficult to check that under our assumptions in Theorem 3.1, one can show the strict convexity of the functional

$$I : \mathbb{H} \to \mathbb{R}, \quad I(\mathbf{u}) = a(\mathbf{u}) - \langle \mathbf{U}, \mathbf{u} \rangle. \tag{3.7}$$

Indeed, let

$$\mathbf{u}_1, \mathbf{u}_0 \in \mathbb{H}, \quad t \in [0,1].$$

We would like to conclude that

$$I(t\mathbf{u}_1 + (1-t)\mathbf{u}_0) \le tI(\mathbf{u}_1) + (1-t)I(\mathbf{u}_0),$$

and equality can only occur if $\mathbf{u}_1 = \mathbf{u}_0$, or $t(1-t) = 0$. Since the second contribution to I is linear in $\mathbf{u}$, it does not perturb its convexity. It suffices to check that indeed

$$a(t\mathbf{u}_1 + (1-t)\mathbf{u}_0) \leq ta(\mathbf{u}_1) + (1-t)a(\mathbf{u}_0),$$

with equality under the same above circumstances. The elementary properties of the bilinear form A lead to

$$ta(\mathbf{u}_1) + (1-t)a(\mathbf{u}_0) - a(t\mathbf{u}_1 + (1-t)\mathbf{u}_0) = t(1-t)a(\mathbf{u}_1 - \mathbf{u}_0).$$

Because the quadratic form a is coercive, it is strictly positive (except for the vanishing vector), and this equality implies the strict convexity of I.

This strict convexity of I, together with its coercivity, has two main consequences that hold true regardless of the dimension of $\mathbb{H}$ (exercise):

1. I admits a unique (global) minimizer;
2. local and global minimizers are exactly the same vectors.

There is another fundamental property that has been used for the quadratic functional I in (3.7). It has been explicitly stated in (3.3). It can also be regarded in a general Banach space. Because it will play a special role for us, we have separated it from our general definition above.

Definition 3.8 Let $I : \mathbb{E} \to \mathbb{R}$ be a functional over a Banach space $\mathbb{E}$. I is said to enjoy the (sequential) weak lower semicontinuity property if

$$I(\mathbf{u}) \leq \liminf_{j\to\infty} I(\mathbf{u}_j)$$

whenever $\mathbf{u}_j \rightharpoonup \mathbf{u}$ in $\mathbb{E}$.

Note that Definition 3.7 introduces lower semicontinuity with respect to convergence in the Banach space itself. We will refer to this, to distinguish it from weak lower semicontinuity, as strong lower semicontinuity whenever appropriate.

Before we focus on convexity, let us describe the direct method of the Calculus of Variations in general terms.

Proposition 3.1 *Let $I : \mathbb{E} \to \mathbb{R}$ be a functional over a reflexive Banach space $\mathbb{E}$, bounded from below by some finite constant, and complying with the two conditions:*

1. *coercivity:*

$$I(\mathbf{u}) \to \infty \text{ when } \|\mathbf{u}\| \to \infty;$$

2. *weak lower semicontinuity according to Definition 3.8.*

Then the there is a global minimizer for I in $\mathbb{E}$.

Proof Let $m \in \mathbb{R}$ be the infimum of I over $\mathbb{E}$, and let $\{\mathbf{u}_j\}$ be minimizing

$$I(\mathbf{u}_j) \searrow m.$$

The coercivity condition implies that $\{\mathbf{u}_j\}$ is uniformly bounded in $\mathbb{E}$, and, hence, by the weak compactness principle Theorem 2.3, there is a subsequence (not relabeled) converging weakly to some $\mathbf{u} \in \mathbb{E}$: $\mathbf{u}_j \rightharpoonup \mathbf{u}$. The weak lower semicontinuity property leads to

$$m \le I(\mathbf{u}) \le \liminf_{j\to\infty} I(\mathbf{u}_j) = m,$$

and $\mathbf{u}$ becomes a global minimizer for I in $\mathbb{E}$. □

This principle is quite easy to understand, and it isolates requirements for a functional to attain its infimum. However, we would like to stress that convexity comes into play concerning the weak lower semicontinuity property.

Proposition 3.2 *Let $I : \mathbb{E} \to \mathbb{R}$ be a convex, (strong) lower semicontinuous functional over a Banach space $\mathbb{E}$. Then it is weak lower semicontinuous.*

Proof The proof is based on one of our main consequences of the Hahn-Banach theorem, specifically Theorem 3.5. We showed there that convex, closed sets in Banach spaces are also weakly closed.

Suppose $\mathbf{u}_j \rightharpoonup \mathbf{u}$, but

$$I(\mathbf{u}) > \liminf_{j\to\infty} I(\mathbf{u}_j).$$

Select $m \in \mathbb{R}$ such that

$$\liminf_{j\to\infty} I(\mathbf{u}_j) < m < I(\mathbf{u}).$$

The sublevel set

$$\mathbf{C} = \{\mathbf{v} \in \mathbb{E} : I(\mathbf{v}) \le m\}$$

is convex and closed, precisely because I is convex and lower semicontinuous. Hence, by Theorem 3.5, $\mathbf{C}$ is also weakly closed. But then the two facts

$$\mathbf{u}_j \rightharpoonup \mathbf{u}, \quad \liminf_{j\to\infty} I(\mathbf{u}_j) < m$$

imply $\mathbf{u} \in \mathbf{C}$, which is impossible if $m < I(\mathbf{u})$. □

As a main consequence, we have an existence abstract theorem for convex functionals that is a direct consequence of Propositions 3.2 and 3.1

Corollary 3.7 *Let $I : \mathbb{E} \to \mathbb{R}$ be a convex, lower semicontinuous functional over a reflexive Banach space $\mathbb{E}$, that is bounded from below by some finite constant, and coercive in the sense*

$$I(\mathbf{u}) \to \infty \text{ when } \|\mathbf{u}\| \to \infty.$$

Then the there is a global minimizer for I in $\mathbb{E}$.

Though this corollary yields a first existence result for a large class of integral functionals, we will focus on an independent, more specific strategy that can be generalized to tackle larger classes of problems, as we will see later in the book, and broader notions of convexity that are central to vector problems.

Conditions to ensure uniqueness of minimizers are, typically, much more restrictive and are usually associated with strict convexity.

Proposition 3.3 *Assume, in addition to hypotheses in Corollary 3.7, that the functional $I : \mathbb{E} \to \mathbb{R}$ is strictly convex in the sense of Definition 3.7. Then there is a unique minimizer for I in $\mathbb{E}$.*

Proof The proof of uniqueness based on strict convexity is quite standard. Suppose there could be two minimizers $\mathbf{u}_1$, $\mathbf{u}_0$,

$$I(\mathbf{u}_1) = I(\mathbf{u}_0) = m = \min I.$$

Then we would have,

$$m \le I((1/2)\mathbf{u}_1 + (1/2)\mathbf{u}_0) \le (1/2)I(\mathbf{u}_1) + (1/2)I(\mathbf{u}_0) = m.$$

Hence

$$I((1/2)\mathbf{u}_1 + (1/2)\mathbf{u}_0) - (1/2)I(\mathbf{u}_1) - (1/2)I(\mathbf{u}_0) = 0$$

and from the strict convexity, we realize that the only alternative is $\mathbf{u}_1 = \mathbf{u}_0$. □

Even for more specific classes of functionals like the ones we will consider in the next chapter, uniqueness is always associated with this last proposition so that strict convexity of functionals need to be enforced. This is however not so for existence, since we can have existence results even though the functional I is not convex.

3.7 Convex Functionals, and the Indirect Method

The direct method of the Calculus of Variations expressed in Proposition 3.1, or more specifically in Corollary 3.7, allows to prove existence of minimizers "directly" without any further computation, except possibly those directed towards checking coercivity and/or convexity, but no calculation is involved in dealing

with the derivative of the functional. In contrast, taking advantage of optimality conditions in the spirit of Lemma 2.7, to decide if there are minimizers, falls under the action of techniques that we could include in what might be called the indirect method.

We already have a very good example to understand this distinction between the direct and the indirect method in the important Lax-Milgram Theorem 3.1 though in this particular, fundamental case both are treated simultaneously. Recall that in this context we are before a symmetric, continuous, coercive bilinear form $A(\mathbf{u}, \mathbf{v})$ over a Hilbert space $\mathbb{H}$, with associated quadratic form

$$a(\mathbf{u}) = \frac{1}{2} A(\mathbf{u}, \mathbf{u}).$$

1. The direct method would ensure, based on the coercivity and the convexity of $a(\mathbf{u})$, that there is a unique minimizer of problem (3.1)

$$\text{Minimize in } \mathbf{u} \in \mathbb{H} : \quad a(\mathbf{u}) - \langle \mathbf{U}, \mathbf{u} \rangle \tag{3.8}$$

   for any given $\mathbf{U} \in \mathbb{H}$.
2. The indirect problem would focus on the condition of optimality (3.2)

$$A(\overline{\mathbf{u}}, \mathbf{v}) = \langle \mathbf{U}, \mathbf{v} \rangle$$

   for every $\mathbf{v} \in \mathbb{H}$, and assuming that we are able to find, independently of the direct method, one solution $\overline{\mathbf{u}}$, argue, again based on the convexity of the quadratic functional in (3.8), that such element $\overline{\mathbf{u}}$ is indeed a minimizer for problem (3.8).

Said differently, the direct method points to a solution, the minimizer, of the conditions of optimality, under smoothness of the functional. As a matter of fact, a minimizer of a variational problem, regardless of how it has been obtained, will be a solution of optimality conditions, under hypotheses guaranteeing smoothness of the functional. From this perspective, we say that existence of solutions of optimality conditions are necessary for the existence of minimizers, even in the absence of convexity. But convexity is required to guarantee that existence of solutions of optimality conditions are sufficient for (global) minimizers.

Consider the variational problem

$$\text{Minimize in } \mathbf{u} \in \mathbb{H} : \quad I(\mathbf{u}) \tag{3.9}$$

where $\mathbb{H}$ is a Hilbert space, and I is differentiable (Definition 2.13).

Proposition 3.4

1. If $\overline{\mathbf{u}} \in \mathbb{H}$ is a minimizer for (3.9), then

$$\langle I'(\overline{\mathbf{u}}), \mathbf{U} \rangle = 0 \tag{3.10}$$

for every $\mathbf{U} \in \mathbb{H}$.

2. *If* $\overline{\mathbf{u}} \in \mathbb{H}$ *is a solution of* (3.10)*, and* I *is convex, then* $\overline{\mathbf{u}}$ *is an optimal solution for* (3.9).
3. *If* I *is strictly convex, then either problem* (3.9) *or* (3.10) *has a unique solution.*

Proof Though the proof is pretty elementary, its significance goes well beyond that simplicity: (3.10) is the abstract expression of the fundamental Euler-Lagrange equation in its weak form. We will review it in more specific frameworks in subsequent chapters.

If $\overline{\mathbf{u}}$ is indeed a minimizer for (3.9), then $\epsilon = 0$ has to be a (global) minimizer for each section

$$\epsilon \in \mathbb{R} \mapsto I(\overline{\mathbf{u}} + \epsilon\mathbf{U}).$$

In particular, its derivative at $\epsilon = 0$ which, by Lemma 2.7, is given by (3.10), must vanish. This does not require any convexity. On the other hand, if I is convex, then

$$I(\mathbf{u} + \mathbf{U}) \geq I(\mathbf{u}) + \langle I'(\mathbf{u}), \mathbf{U}\rangle$$

for every $\mathbf{u}$, $\mathbf{U}$ in $\mathbb{H}$ (exercise below). In particular, if $\mathbf{u} = \overline{\mathbf{u}}$ complies with (3.10), then we immediately see that

$$I(\overline{\mathbf{u}} + \mathbf{U}) \geq I(\overline{\mathbf{u}})$$

for every $\mathbf{U}$, and $\overline{\mathbf{u}}$ becomes a minimizer. The uniqueness has already been treated in the last section under strict convexity. This is left as an exercise. □

The relevance of the simple ideas in this section will be better appreciated in some cases when the direct method for integral functionals is inoperative because of lack of appropriate coercivity properties. In such a situation, the indirect method may be explored to see if existence of global minimizers may be reestablished.

3.8 Stampacchia's Theorem: Variational Inequalities

Stampacchia's theorem is a fundamental statement to deal with quadratic, not necessarily symmetric, forms in a Hilbert space $\mathbb{H}$. It is the simplest example of a variational inequality, and it is typically used to have the Lax-Milgram lemma as an easy corollary (Exercise 10). As we have shown above, we have followed a different, independent, more variational, route to prove that crucial lemma. However, it is important to know this other fundamental result by Stampacchia. Its proof is a beautiful application of the classical contraction principle Theorem 1.1.

Theorem 3.6 *Let* $A(\mathbf{u}, \mathbf{v})$ *be a continuous, coercive bilinear form in a Hilbert space* $\mathbb{H}$*, and let* $\mathbf{K} \subset \mathbb{H}$ *be non-empty, closed, and convex. For every* $T \in \mathbb{H}'$,

there is a unique $\mathbf{u} \in \mathbf{K}$ *such that*

$$A(\mathbf{u}, \mathbf{v} - \mathbf{u}) \geq \langle T, \mathbf{v} - \mathbf{u} \rangle$$

for every $\mathbf{v} \in \mathbf{K}$. *In, in addition, A is symmetric, the vector* $\mathbf{u}$ *is characterized as the unique minimizer of the quadratic functional*

$$\mathbf{v} \mapsto \frac{1}{2} A(\mathbf{v}, \mathbf{v}) - \langle T, \mathbf{v} \rangle$$

over $\mathbf{K}$.

Proof Through the Riesz-Fréchet representation theorem Proposition 2.14, it is easy to show the existence of a linear, continuous operator $\mathbf{A} : \mathbb{H} \to \mathbb{H}$ such that

$$A(\mathbf{u}, \mathbf{v}) = \langle \mathbf{A}\mathbf{u}, \mathbf{v} \rangle, \quad \|\mathbf{A}\mathbf{u}\| \leq C\|\mathbf{u}\|, \quad \mathbf{u}, \mathbf{v} \in \mathbb{H}, \tag{3.11}$$

and, moreover,

$$\langle \mathbf{A}\mathbf{u}, \mathbf{u} \rangle \geq \frac{1}{C}\|\mathbf{u}\|^2, \quad \mathbf{u} \in \mathbb{H}, \tag{3.12}$$

for some positive constant C. Similarly, through the same procedure, there is some $\mathbf{t} \in \mathbb{H}$ with

$$\langle T, \mathbf{v} \rangle = \langle \mathbf{t}, \mathbf{v} \rangle, \quad \mathbf{v} \in \mathbb{H}.$$

Note that the left-hand side is the duality pair in the Hilbert space $\mathbb{H}$, while the right-hand side is the inner product in $\mathbb{H}$. There is, we hope, no confusion in using the same notation. In these new terms, we are searching for a vector $\mathbf{u} \in \mathbb{H}$ such that

$$\langle \mathbf{A}\mathbf{u}, \mathbf{v} - \mathbf{u} \rangle \geq \langle \mathbf{t}, \mathbf{v} - \mathbf{u} \rangle, \quad \mathbf{v} \in \mathbb{H}. \tag{3.13}$$

If for a positive r, we recast (3.13) in the form

$$\langle r\mathbf{t} - r\mathbf{A}\mathbf{u} + \mathbf{u} - \mathbf{u}, \mathbf{v} - \mathbf{u} \rangle \leq 0, \quad \mathbf{v} \in \mathbb{H},$$

we realize, after Proposition 2.11 concerned with the orthogonal projection onto a convex set, that the previous inequality is equivalent to the functional equation

$$\mathbf{u} = \pi_{\mathbf{K}}(r\mathbf{t} - r\mathbf{A}\mathbf{u} + \mathbf{u}).$$

We would like to interpret this equation as a fixed point for the mapping

$$\mathbf{u} \mapsto \mathbf{T}\mathbf{u} \equiv \pi_{\mathbf{K}}(r\mathbf{t} - r\mathbf{A}\mathbf{u} + \mathbf{u}).$$

The contraction principle Theorem 1.1 guarantees the existence of a unique fixed point for such a mapping, provided that it is a contraction. Since the norm of every projection is less than unity (check the discussion after the proof of Proposition 2.11), we find that

$$\begin{aligned}\|\mathbf{T}\mathbf{u}_1 - \mathbf{T}\mathbf{u}_2\|^2 \leq& \|\mathbf{u}_1 - \mathbf{u}_2 - r\mathbf{A}(\mathbf{u}_1 - \mathbf{u}_2)\|^2 \\ =& \|\mathbf{u}_1 - \mathbf{u}_2\|^2 - 2r\langle \mathbf{u}_1 - \mathbf{u}_2, \mathbf{A}(\mathbf{u}_1 - \mathbf{u}_2)\rangle + r^2\|\mathbf{A}(\mathbf{u}_1 - \mathbf{u}_2)\|^2 \\ \leq& \|\mathbf{u}_1 - \mathbf{u}_2\|^2 - 2r\frac{1}{C^2}\|\mathbf{u}_1 - \mathbf{u}_2\|^2 + C^2r^2\|\mathbf{u}_1 - \mathbf{u}_2\|^2 \\ =& \left(1 - \frac{2r}{C^2} + r^2C^2\right)\|\mathbf{u}_1 - \mathbf{u}_2\|^2,\end{aligned}$$

by (3.11) and (3.12). If we select

$$0 < r < \frac{2}{C^4},$$

we clearly see that $\mathbf{T}$ becomes indeed a contraction, and it admits a unique fixed point $\mathbf{u}$ which is the vector sought as remarked earlier.

The symmetric case and the minimization property is left as an exercise, though it takes us back to the Lax-Milgram theorem. □

3.9 Exercises

1. Prove the third part of Corollary 3.2 relying on the other two parts.
2. For a finite set of vectors

$$\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\} \subset \mathbb{E}$$

in a normed space $\mathbb{E}$, show that the three numbers

$$\sup\{\|\sum_j \lambda_j \mathbf{x}_j\| : \lambda_j = \pm 1\},$$

$$\sup\{\|\sum_j \lambda_j \mathbf{x}_j\| : -1 \leq \lambda_j \leq 1\},$$

$$\sup\{\sum_j |\langle \mathbf{x}', \mathbf{x}_j\rangle| : \mathbf{x}' \in \mathbb{E}', \|\mathbf{x}'\| \leq 1\},$$

are identical.
3. Show that a difference set $\mathbf{C} = \mathbf{F} - \mathbf{G}$ is closed if $\mathbf{F}$ is closed and $\mathbf{G}$, compact.

4. Let

$$f(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R},$$

be coercive and strictly convex. Prove that f admits a unique local minima which is also global.

5. (a) Let a function

$$f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$$

be C^1. Show that it is convex if and only if

$$f(\mathbf{x}+\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot \mathbf{y}$$

for arbitrary $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(b) Suppose the functional $I : \mathbb{E} \to \mathbb{R}$ is differentiable and convex, where $\mathbb{E}$ is a Hilbert space. Prove that

$$I(\mathbf{u}+\mathbf{U}) \geq I(\mathbf{u}) + \langle I'(\mathbf{u}), \mathbf{U}\rangle$$

for every $\mathbf{u}$, $\mathbf{U}$ in $\mathbb{E}$. If, in addition, I is strictly convex then the equality

$$I(\mathbf{u}+\mathbf{U}) = I(\mathbf{u}) + \langle I'(\mathbf{u}), \mathbf{U}\rangle$$

is only valid if $\mathbf{U} \equiv \mathbf{0}$. Based on this, prove the third item of Proposition 3.4.

6. Show that if u is a smooth solution of (3.5), then (3.4) is correct for every $v \in H_0^1(J)$.

7. Consider the Hilbert space $\mathbb{H} = H_0^1(0, 1)$.

(a) Argue that

$$\langle u, v\rangle = \int_0^1 u'(t)v'(t)\,dt$$

is an equivalent inner-product in $\mathbb{H}$, and regard $\mathbb{H}$ as endowed with it.

(b) Consider the bilinear form

$$A(u, v) = \int_0^1 tu'(t)v'(t)\,dt,$$

and study the corresponding Lax-Milgram problem (3.1) for $\mathbf{U} = 0$. Show that there is no minimizer for it. What is the explanation for such a situation?

8. In the context of the Lax-Milgram theorem, consider the map $\mathbf{T} : \mathbb{H} \to \mathbb{H}$ taking each $\mathbf{U} \in \mathbb{H}$ into the unique solution $\overline{\mathbf{u}}$. Show that this is a linear continuous operator.
9. Show Stampacchia's theorem for the symmetric case by defining on $\mathbb{H}$ a new, equivalent inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle_A = A(\mathbf{u}, \mathbf{v}),$$

and using the orthogonal projection theorem for it. In this same vein, argue that the solution $\mathbf{u}$ is characterized as the unique minimizer of the corresponding quadratic functional in the statement of Stampachhia's theorem.
10. Prove the Lax-Milgram lemma from Stampachia's theorem.
11. For $f(x) \in L^2(J)$, define the functional $I : H_0^1(J) \to \mathbb{R}$ through

$$I(u) = \int_J F(v(x))\,dx,$$

$$-(u'(x) + v'(x))' + f(x) = 0 \text{ in } J, \quad v \in H_0^1(J).$$

Explore conditions on the integrand $F : \mathbb{R} \to \mathbb{R}$ allowing for the direct method to be applied.
12. For an open, convex subset D in $\mathbb{R}^N$, let $u(\boldsymbol{x}) : D \to \mathbb{R}$ be a convex function. Let $(\mathbb{X}, \mu)$ be a probability space, and

$$v_j(\mathbf{x}) : \mathbb{X} \to \mathbb{R}, \quad 1 \le j \le N,$$

integrable functions such that the joint mapping

$$\mathbf{x} \mapsto (v_j(\mathbf{x}))_j \in D$$

and the composition

$$\mathbf{x} \mapsto u(v_1(\mathbf{x}), v_2(\mathbf{x}), \dots, v_N(\mathbf{x}))$$

is integrable (with respect to μ).

(a) Argue that

$$\left(\int_{\mathbb{X}} v_j(\mathbf{x})\,d\mu(\mathbf{x})\right)_j \in D.$$

(b) Prove Jensen's inequality

$$u\left(\int_{\mathbb{X}} v_1(\mathbf{x})\,d\mu(\mathbf{x}), \dots, \int_{\mathbb{X}} v_N(\mathbf{x})\,d\mu(\mathbf{x})\right) \leq \int_{\mathbb{X}} u(v_1(\mathbf{x}), \dots, v_N(\mathbf{x}))\,d\mu(\mathbf{x}).$$

(c) Figure out a suitable choice of N, D, u, $\mathbb{X}$ and μ, to prove that the geometric mean of m positive numbers r_j is always smaller than the arithmetic mean

$$\Pi_j r_j^{\lambda_j} \leq \sum_j \lambda_j r_j, \quad \lambda_j > 0, \sum_j \lambda_j = 1.$$

13. Explore conditions on the integrand

$$F(u, v) : \mathbb{R}^2 \to \mathbb{R}$$

to guarantee that the functional

$$I(u) = \int_J \int_J f(u(x), u(y))\,dx\,dy$$

is convex. J is an interval in $\mathbb{R}$.

14. (a) Consider a functional of the form

$$I(u) = \int_0^1 \Phi\left(\int_0^1 W(x, y, u(y))\,dy\right)\,dx$$

for certain functions

$$W(x, y, u, v) : (0, 1)^2 \times \mathbb{R} \to \mathbb{R}, \quad \Phi(t) : \mathbb{R} \to \mathbb{R}.$$

Derive a set of sufficient conditions for the functional I to be weak lower semicontinuous.

(b) The generalization to a funcional of the form

$$I(u) = \int_0^1 \Phi\left(\int_0^1 W(x, y, u(x), u(y))\,dy\right)\,dx$$

is much more difficult. Try to argue why.

15. Let functions

$$f_i(x, u) : (0, 1) \times \mathbb{R} \to \mathbb{R}, \quad 1 \leq i \leq n, \quad \boldsymbol{f}(x, u) = (f_i(x, u))_i,$$

be given, and $F : \mathbb{R}^n \to \mathbb{R}$. Consider the functional

$$I(u) = F\left(\int_J f(x, u(x))\,dx\right).$$

Explore both the direct and indirect methods for this family of functionals.

16. Let $\{u_i\}$ be an orthonormal basis of $L^2(\Omega)$ for a open set $\Omega \subset \mathbb{R}^N$, and consider

$$s_N : L^2(\Omega) \to \mathbb{R},$$

$$N s_N(f)^2 = \int_{\Omega\times\Omega} f(\boldsymbol{x}) \sum_{i=1}^{N} u_i(\boldsymbol{x}) u_i(\boldsymbol{y})\, f(\boldsymbol{y})\, d\boldsymbol{x}\, d\boldsymbol{y}.$$

   (a) Check that each s_N is a semi-norm in $L^2(\Omega)$.
   (b) Show that the usual topology of $L^2(\Omega)$ (the one associated with its usual inner product) is the coarser one that makes the collection $\{s_N\}_N$ continuous.

17. For the functional

$$E : H_0^1(0,1) \to \mathbb{R}, \quad E(u) = \int_0^1 [\psi(u'(x)) + \phi(u(x))]\,dx$$

where ψ and ϕ are C^1-, real functions such that

$$|\psi(u)|, |\phi(u)| \le C(1 + u^2)$$

for a positive constant C, calculate, as explicitly as possible, its derivative $E'(u)$ for an arbitrary $u \in H_0^1(0,1)$.

Chapter 4
The Calculus of Variations for One-dimensional Problems

4.1 Overview

The goal of this chapter is to study one-dimensional variational problems where one tries to minimize an integral functional of the form

$$I(\mathbf{u}) = \int_J F(x, \mathbf{u}(x), \mathbf{u}'(x))\, dx \tag{4.1}$$

among a certain class of functions (paths) $\mathcal{A}$. Here J is a finite interval in $\mathbb{R}$, and the integrand

$$F(x, \mathbf{u}, \mathbf{v}) : J \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

is a function whose properties are the main object of our concern. Fundamental ingredients to be examined are function spaces where competing paths

$$\mathbf{u}(x) : J \to \mathbb{R}^n$$

are to be taken from. Additional constraints, typically in the form of prescribed end-point conditions, are also to be taken into account. All of these various requirements will be hauled into the collection of competing paths $\mathcal{A}$.

Our goal is to understand these two principal issues for functionals of this integral type:

1. isolate properties on the integrand F to ensure the existence of at least one minimizer in a natural function space, under typical sets of constraints; and
2. determine additional properties, the so-called optimality conditions, that such minimizers should comply with precisely because they are optimal solutions of a given variational problem.

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4_4

As a general rule, existence theorems for minimizers of functionals can be established in Banach spaces; however, optimality conditions involving derivatives require, at least in a first step, an inner product, and so we will look at them in a Hilbert space scenario. Optimality conditions in Banach spaces demand additional training that we will overlook here.

Though variants can, and should, be accepted and studied, our basic model problem will be

$$\text{Minimize in } \mathbf{u} \in \mathcal{A}: \quad I(\mathbf{u})$$

where $I(\mathbf{u})$ is given in (4.1), and $\mathcal{A}$ is a subset of a suitable Sobolev space usually determined through fixed, preassigned values at the end-points of J

$$\mathbf{u}(x_0) = \mathbf{u}_0, \quad \mathbf{u}(x_1) = \mathbf{u}_1, \quad J = (x_0, x_1) \subset \mathbb{R}, \mathbf{u}_0, \mathbf{u}_1 \in \mathbb{R}^n.$$

There are many more topics one can look at even for one-dimensional problems: important issues hardly end with this introductory chapter. As usual, we will indicate some of these in the final Appendix of the book.

Since convexity will play a central role in this book henceforth, we will remind readers of the main inequality involving convex functions: Jensen's inequality. It has already been examined in Exercise 12 of Chap. 3. Convexity for functionals has already been treated in the same chapter. We will reserve a more practical look at convexity until Chap. 8 where functions and integrands depend on several variables. Though this can be the case for vector, uni-dimensional problems too, we defer a more systematic discussion until then. Most of the interesting examples examined in this chapter correspond to uni-dimensional, scalar problems and so convexity is mainly required for real functions of one variable.

4.2 Convexity

Though often convexity is defined for functions taking on the value $+\infty$, and this is quite efficient, we will only consider real functions defined on subsets.

Definition 4.1 A function

$$\phi : D \subset \mathbb{R}^N \to \mathbb{R}$$

is convex if D is a convex set, and

$$\phi(t_1\mathbf{x}_1 + t_2\mathbf{x}_2) \le t_1\phi(\mathbf{x}_1) + t_2\phi(\mathbf{x}_2), \quad \mathbf{x}_1, \mathbf{x}_2 \in D, t_1, t_2 \ge 0, t_1 + t_2 = 1. \tag{4.2}$$

If there is no mention about the set D, one consider it to be the whole space $\mathbb{R}^N$ or the natural domain of definition of ϕ (which must be a convex set). Moreover, such

a function ϕ is said to be strictly convex if the equality in (4.2)

$$\phi(t_1\mathbf{x}_1 + t_2\mathbf{x}_2) = t_1\phi(\mathbf{x}_1) + t_2\phi(\mathbf{x}_2), \quad \mathbf{x}_1, \mathbf{x}_2 \in D, t_1, t_2 \geq 0, t_1 + t_2 = 1,$$

can only occur when either $\mathbf{x}_1 = \mathbf{x}_2$ or $t_1 t_2 = 0$.

Though we will be more explicit in Chap. 8 when dealing with high-dimensional variational problems, we anticipate two basic facts which are easily checked:

1. Every linear (affine) function is always convex, though not strictly convex.
2. The supremum of convex functions is also convex.

One main consequence is that the supremum of linear functions is always convex. In fact, this is always the case. To dig deeper into these issues, we introduce the following concept.

Definition 4.2 For a function $\phi(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}$, we define

$$\mathbf{C}_L\phi = \sup\{\psi : \psi \leq \phi, \psi, \text{ linear}\},$$
$$\mathbf{C}_C\phi = \sup\{\psi : \psi \leq \phi, \psi, \text{ convex}\}.$$

It is obvious that

$$\mathbf{C}_L\phi \leq \mathbf{C}_C\phi \leq \phi,$$

and that if ϕ is convex, then $\mathbf{C}_C\phi = \phi$. The following important fact is an unexpected consequence of Theorems 3.3 or 3.4. In the finite-dimensional case, both are equivalent.

Proposition 4.1 *A function $\phi(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}$ is convex if and only if*

$$\phi \equiv \mathbf{C}_L\phi \equiv \mathbf{C}_C\phi.$$

Proof Suppose ϕ is a convex function. Its epigraph

$$E = \{(\mathbf{x}, t) \in \mathbb{R}^{N+1} : t \geq \phi(\mathbf{x})\}$$

is, then, a convex set. This is straightforward to check. If we assume that there is a point $\boldsymbol{y}$ where $\mathbf{C}_L\phi(\boldsymbol{y}) < \phi(\boldsymbol{y})$, then the two convex sets

$$E, \quad \{(\mathbf{y}, \mathbf{C}_L\phi(\mathbf{y}))\}$$

are disjoint with the first, closed, and the second, compact. Theorem 3.4 guarantees that there is a closed hyperplane in $\mathbb{R}^{N+1}$ separating them. Hyperplanes in $\mathbb{R}^{N+1}$ are of the form

$$\mathbf{u} \cdot \mathbf{x} + u x_{N+1} = c, \quad \mathbf{u}, \mathbf{x} \in \mathbb{R}^N, u, c \in \mathbb{R};$$

and hence there is such $\mathbf{u} \in \mathbb{R}^N$, $u \in \mathbb{R}$, with

$$\begin{aligned} \mathbf{u} \cdot \mathbf{x} + ut &> c \text{ when } t \geq \phi(\mathbf{x}), \\ \mathbf{u} \cdot \mathbf{y} + u\mathbf{C}_L\phi(\mathbf{y}) &< c. \end{aligned} \tag{4.3}$$

We claim that u cannot be zero, for if it were, we would have

$$\mathbf{u} \cdot \mathbf{x} < c < \mathbf{u} \cdot \mathbf{y}$$

with no restriction on $\mathbf{x} \in \mathbb{R}^N$. This is clearly impossible. Suppose that, without loss of generality, u is positive, and define

$$L\mathbf{x} = -\frac{1}{u}\mathbf{u} \cdot \mathbf{x} + \frac{1}{u}c.$$

then (4.3) for $t = \phi(\boldsymbol{x})$ would lead to $L\mathbf{x} < \phi(\mathbf{x})$ for every $\mathbf{x}$, and yet $L\mathbf{y} > \mathbf{C}_L\phi(\mathbf{y})$, which is a contradiction with the definition of $\mathbf{C}_L\phi$. □

The most fundamental fact involving convexity is Jensen's inequality. It is almost an immediate consequence of Proposition (4.1).

Theorem 4.1 *If $\phi : \mathbb{R}^N \to \mathbb{R}$ is a convex function, and μ is a probability measure supported in $\mathbb{R}^N$, then*

$$\phi\left(\int_{\mathbb{R}^N} \mathbf{x}\, d\mu(\mathbf{x})\right) \leq \int_{\mathbb{R}^N} \phi(\mathbf{x})\, d\mu(\mathbf{x}).$$

If ϕ is strictly convex, and

$$\int_{\mathbb{R}^N} \phi(\mathbf{x})\, d\mu(\mathbf{x}) = \phi\left(\int_{\mathbb{R}^N} \mathbf{x}\, d\mu(\mathbf{x})\right),$$

then μ is a Dirac mass $\delta_{\mathbf{z}}$ supported at

$$\mathbf{z} = \int_{\mathbb{R}^N} \mathbf{x}\, d\mu(\mathbf{x}).$$

Proof The fact that μ is a positive measure with total mass 1, implies that if L is a linear function such that $L \leq \phi$, then because integration is a linear operation,

$$\begin{aligned} L\left(\int_{\mathbb{R}^N} \boldsymbol{x}\, d\mu(\boldsymbol{x})\right) &= \int_{\mathbb{R}^N} L(\boldsymbol{x})\, d\mu(\boldsymbol{x}) \\ &\leq \int_{\mathbb{R}^N} \phi(\boldsymbol{x})\, d\mu(\boldsymbol{x}). \end{aligned}$$

If we take the supremum on the left-hand side among all linear functions with $L \leq \phi$, by Proposition 4.1, we have our inequality.

If, in fact, the last inequality turns out to be an equality, then the support of μ must be contained where ϕ equals $\mathbf{C}_L\phi$; but the strict inequality of ϕ implies that the support of μ must be a singleton. □

It is not difficult to generalize in a suitable way these concepts and results for functions that can take on the value $+\infty$, by restricting all operations to the domain of ϕ.

4.3 Weak Lower Semicontinuity for Integral Functionals

So far, we have been exposed to abstract, general functionals defined over a Banach or Hilbert space, and have anticipated the central role played by the convexity property of the functional under consideration. When one restricts attention to special classes of functionals, finer results can typically be shown. We start now the study of integral functionals of the type

$$I(\mathbf{u}) = \int_J F(x, \mathbf{u}(x), \mathbf{u}'(x))\, dx \tag{4.4}$$

defined over suitable subsets of Sobolev spaces of competing functions or paths $\mathbf{u}$. Our model problem will insist on end-point conditions of the form

$$\mathbf{u}(x_0) = \mathbf{u}_0, \quad \mathbf{u}(x_1) = \mathbf{u}_1$$

if $J = (x_0, x_1)$ is a finite interval in $\mathbb{R}$, and $\mathbf{u}_0, \mathbf{u}_1 \in \mathbb{R}^n$, though other restrictions can be adapted to variational methods. The properties of the integrand

$$F(x, \mathbf{u}, \mathbf{U}) : J \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \tag{4.5}$$

will be crucial to our results. In particular, we are mostly interested in understanding for which such densities the direct method Proposition 3.1 can be implemented to produce existence results for minimizers for I. One way or another the convexity of I or of F will have to play a central role.

To be more specific, and for the sake of definiteness, once J is given, take a finite $p > 1$, and let

$$\mathbb{E} = W^{1,p}(J; \mathbb{R}^n)$$

be our basic, underlying Banach space where we will consider our variational problems. For $p \neq 2$, it is a Banach space, while for $p = 2$, $\mathbb{E}$ becomes

$\mathbb{H} = H^1(J; \mathbb{R}^n)$, a Hilbert space. Our model problem will be

$$\text{Minimize in } \mathbf{u} \in \mathbb{E} = W^{1,p}(J; \mathbb{R}^n): \quad I(\mathbf{u}) = \int_J F(x, \mathbf{u}(x), \mathbf{u}'(x))\, dx$$

subject to

$$\mathbf{u}(x_0) = \mathbf{u}_0, \quad \mathbf{u}(x_1) = \mathbf{u}_1.$$

Strictly speaking, the class

$$\mathcal{A} = \{\mathbf{u} \in W^{1,p}(J; \mathbb{R}^n) : \mathbf{u}(x_0) = \mathbf{u}_0, \mathbf{u}(x_1) = \mathbf{u}_1\},$$

is not a vector space. However, if we set

$$L : J \to \mathbb{E}, \quad L(x) = \frac{x - x_0}{x_1 - x_0}\mathbf{u}_1 + \frac{x_1 - x}{x_1 - x_0}\mathbf{u}_0,$$

then

$$\mathcal{A} = L + W_0^{1,p}(J; \mathbb{R}^n),$$
$$W_0^{1,p}(J; \mathbb{R}^n) = \{\mathbf{u} \in W^{1,p}(J; \mathbb{R}^n) : \mathbf{u}(x_0) = \mathbf{u}(x_1) = \mathbf{0}\},$$

and $W_0^{1,p}(J; \mathbb{R}^n)$ is a vector subspace of $\mathbb{E}$. Competing paths for our model problem can then be written in the form

$$\mathbf{u} = L + \overline{\mathbf{u}}, \quad \overline{\mathbf{u}} \in W_0^{1,p}(J; \mathbb{R}^n).$$

We will not insist in this point any longer, and, without further notice, we will work with the class $\mathcal{A}$ as if it were a Banach subspace of $\mathbb{E}$.

The direct method Proposition 3.1 will guide us. In fact, we could already state a first result through Proposition 4.1. But since we are after a more general existence result which is specific for our integral functionals and, hence, it can guide us in other situations, we will follow a different strategy.

Let us focus on functional (4.4) for an integrand as in (4.5). $F(x, \mathbf{u}, \mathbf{U})$ is assumed to be continuous on pairs $(\mathbf{u}, \mathbf{U})$, but only measurability on the spatial variable x is assumed. It is clear that if

$$F(x, \cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

is convex for each $x \in J$, then the corresponding functional I in (4.4) is also convex. In this way, and under additional suitable assumptions, we could rely on Proposition 4.1, as already indicated, for weak lower semicontinuity, and eventually have an existence result through Corollary 3.7. However, our emphasis is placed in

the fact that only convexity of $F(x, \mathbf{u}, \mathbf{U})$ with respect to variable $\mathbf{U}$ is necessary for our functional I to be weakly lower semicontinuous in Sobolev spaces. Our main result on weak lower semicontinuity follows.

Theorem 4.2 *Let*

$$F(x, \mathbf{u}, \mathbf{U}) : J \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

be continuous in $(\mathbf{u}, \mathbf{U})$ *for a.e.* $x \in J$*, and measurable in* x *for every pair* $(\mathbf{u}, \mathbf{U})$*. The associated functional* I *in* (4.4) *is weak lower semicontinuous in* $W^{1,p}(J, \mathbb{R}^n)$ *if and only if* $F(x, \mathbf{u}, \cdot)$ *is convex for a.e.* $x \in J$ *and every* $\mathbf{u} \in \mathbb{R}^n$*.*

The proof of this main result is our goal in the rest of this section.

The most interesting part for us is the sufficiency. The necessity part is, however, relevant too because it is informing us that there cannot be surprises concerning convexity: weak lower semicontinuous functional cannot escape convexity. This is in deep contrast with the vector case in which one considers competing fields

$$\mathbf{u} \in W^{1,p}(\Omega, \mathbb{R}^n), \quad \Omega \subset \mathbb{R}^N,$$

and both dimensions n and N are at least two. This requires, as a preliminary step, a serious concern about high-dimensional Sobolev spaces. We will treat these in the third part of the book. For vector problems, the necessity part of Theorem 4.2 dramatically fails, and thus opens the door to more general functionals than those having a convex dependence on $\mathbf{U}$.

We begin by arguing the necessity part. Suppose the functional I in (4.4) is weak lower semicontinuous, and let

$$a \in J, \quad \mathbf{u}_0, \mathbf{U}_0, \mathbf{U}_1 \in \mathbb{R}^n. \tag{4.6}$$

Take $\delta > 0$ sufficiently small so that

$$J_{a,\delta} = [a - \delta/2, a + \delta/2] \subset J = (x_0, x_1).$$

For each j, divide $J_{a,\delta}$ in $2j$ equal subintervals of length $\delta/(2j)$, and let $\chi_j(x)$ be the characteristic function of the odd subintervals, so that $1 - \chi_j(x)$, restricted to $J_{a,\delta}$ is the characteristic function of the even subintervals of such family. Define

$$\mathbf{u}_j(x) = \begin{cases} \mathbf{u}_0, & x \in (x_0, a - \delta/2) \\ \mathbf{u}_0 + \int_{a-\delta/2}^{x} \chi_j(s)\, ds \mathbf{U}_1 + \int_{a-\delta/2}^{x} (1 - \chi_j(s))\, ds \mathbf{U}_0, & x \in J_{a,\delta}, \\ \mathbf{u}_0 + (\delta/2)(\mathbf{U}_1 + \mathbf{U}_0), & x \in (a + \delta/2, x_1). \end{cases}$$

We claim that each $\mathbf{u}_j \in W^{1,p}(J;\mathbb{R}^n)$. In fact, it is easy to check that it is continuous as it is continuous in each subinterval

$$(x_0, a-\delta/2), \quad J_{a,\delta}, \quad (a+\delta/2, x_1),$$

separately, and it matches across common end-points. On the other hand, the derivative $\mathbf{u}_j'(x)$ turns out to be, except in a finite number of points,

$$\mathbf{u}_j'(x) = \begin{cases} \mathbf{0}, & x \in (x_0, a-\delta/2) \\ \chi_j(x)\mathbf{U}_1 + (1-\chi_j(x))\mathbf{U}_0, & x \in J_{a,\delta}, \\ \mathbf{0}, & x \in (a+\delta/2, x_1), \end{cases}$$

and this is definitely a path in $L^p(J;\mathbb{R}^n)$. In addition, if we recall Example 2.12, we conclude that $\chi_j \rightharpoonup 1/2$ in $J_{a,\delta}$, and $\mathbf{u}_j' \rightharpoonup \mathbf{U}$ where

$$\mathbf{U}(x) = \begin{cases} \mathbf{0}, & x \in (x_0, a-\delta/2) \\ (1/2)\mathbf{U}_1 + (1/2)\mathbf{U}_0, & x \in J_{a,\delta}, \\ \mathbf{0}, & x \in (a+\delta/2, x_1). \end{cases}$$

If, finally, we define

$$\mathbf{u}(x) = \mathbf{u}_0 + \int_{x_0}^{x} \mathbf{U}(s)\,ds,$$

then, by Proposition 2.8, we can conclude that

$$\mathbf{u}_j \rightharpoonup \mathbf{u} \text{ in } W^{1,p}(J;\mathbb{R}^n).$$

If our functional I in (4.4) is truly weak lower semicontinuous, we ought to have

$$I(\mathbf{u}) \le \liminf_{j\to\infty} I(\mathbf{u}_j). \tag{4.7}$$

Let us examine the two sides of this inequality. For the left-hand side, we find

$$\begin{aligned} I(\mathbf{u}) = &\int_{x_0}^{a-\delta/2} F(x, \mathbf{u}_0, \mathbf{0})\,dx \\ &+ \int_{a-\delta/2}^{a+\delta/2} F(x, \mathbf{u}_0 + (x/2)(\mathbf{U}_1 + \mathbf{U}_0), (1/2)(\mathbf{U}_1 + \mathbf{U}_0))\,dx \\ &+ \int_{a+\delta/2}^{x_1} F(x, \mathbf{u}_0 + (\delta/2)(\mathbf{U}_1 + \mathbf{U}_0), \mathbf{0})\,dx. \end{aligned}$$

Similarly,

$$
\begin{aligned}
I(\mathbf{u}_j) = &\int_{x_0}^{a-\delta/2} F(x, \mathbf{u}_0, \mathbf{0})\, dx \\
&+ \int_{a-\delta/2}^{a+\delta/2} F(x, \mathbf{u}_j(x), \chi_j(x)\mathbf{U}_1 + (1-\chi_j(x))\mathbf{U}_0)\, dx \\
&+ \int_{a+\delta/2}^{x_1} F(x, \mathbf{u}_0 + (\delta/2)(\mathbf{U}_1 + \mathbf{U}_0), \mathbf{0})\, dx.
\end{aligned}
$$

Concerning the central contribution, it is reasonable to expect (proposed as an exercise) that the difference

$$
\begin{aligned}
&\int_{a-\delta/2}^{a+\delta/2} F(x, \mathbf{u}_j(x), \chi_j(x)\mathbf{U}_1 + (1-\chi_j(x))\mathbf{U}_0)\, dx \\
-&\int_{a-\delta/2}^{a+\delta/2} F(x, \mathbf{u}_0 + (x/2)(\mathbf{U}_1 + \mathbf{U}_0), \chi_j(x)\mathbf{U}_1 + (1-\chi_j(x))\mathbf{U}_0)\, dx
\end{aligned}
$$

tends to zero as $j \to \infty$ because the convergence $\mathbf{u}_j \to \mathbf{u}$ takes place in $L^\infty(J; \mathbb{R}^n)$, and F is continuous with respect to the $\mathbf{u}$-variable. Since the first and third terms of the above decompositions are identical for both sides in inequality (4.7), such inequality and the preceding remarks lead to

$$
\begin{aligned}
&\int_{a-\delta/2}^{a+\delta/2} F(x, \mathbf{u}_0 + (x/2)(\mathbf{U}_1 + \mathbf{U}_0), (1/2)(\mathbf{U}_1 + \mathbf{U}_0))\, dx \\
\le \liminf_{j\to\infty} &\int_{a-\delta/2}^{a+\delta/2} F(x, \mathbf{u}_0 + (x/2)(\mathbf{U}_1 + \mathbf{U}_0), \chi_j(x)\mathbf{U}_1 + (1-\chi_j(x))\mathbf{U}_0)\, dx.
\end{aligned}
$$

But the integral in the limit of the right-hand side of this inequality can be broken into two parts

$$
\begin{aligned}
&\int_{J_{a,\delta}\cap\{\chi_j=1\}} F(x, \mathbf{u}_0 + (x/2)(\mathbf{U}_1 + \mathbf{U}_0), \mathbf{U}_1)\, dx \\
+&\int_{J_{a,\delta}\cap\{\chi_j=0\}} F(x, \mathbf{u}_0 + (x/2)(\mathbf{U}_1 + \mathbf{U}_0), \mathbf{U}_0)\, dx,
\end{aligned}
$$

For δ sufficiently small, each of these two integrals essentially are

$$
\frac{1}{2} F(a, \mathbf{u}_0, \mathbf{U}_1) + \frac{1}{2} F(a, \mathbf{u}_0, \mathbf{U}_0),
$$

and the weak lower semicontinuity inequality becomes

$$F(a, \mathbf{u}_0, (1/2)(\mathbf{U}_1 + \mathbf{U}_0)) \le \frac{1}{2}F(a, \mathbf{u}_0, \mathbf{U}_1) + \frac{1}{2}F(a, \mathbf{u}_0, \mathbf{U}_0).$$

The arbitrariness in (4.6) leads to the claimed convexity. There is however a few technical steps to be covered for a full proof that are left as exercises.

For the sufficiency part, our main tool is the classical Jensen's inequality, Theorem 4.1. It was also treated in Exercise 12 of Chap. 3. Turning back to the sufficiency part of Theorem 4.2, suppose

$$\mathbf{u}_j \rightharpoonup \mathbf{u} \text{ in } W^{1,p}(J; \mathbb{R}^n). \tag{4.8}$$

For a positive integer l, write

$$\int_J F(x, \mathbf{u}(x), \mathbf{u}'(x))\, dx = \sum_{i=1}^{l} \int_{J_i} F(x, \mathbf{u}(x), \mathbf{u}'(x))\, dx$$

where the full interval of integration J is the disjoint union of the J_i's and the measure of each J_i is $1/l$. Then

$$\int_J F(x, \mathbf{u}(x), \mathbf{u}'(x))\, dx \sim \sum_{i=1}^{l} \frac{1}{l} F\left(l \int_{J_i} x\, dx, l \int_{J_i} \mathbf{u}(x)\, dx, l \int_{J_i} \mathbf{u}'(x)\, dx\right).$$

Because of the weak convergence (4.8), and the continuity of F,

$$\begin{aligned}
&\lim_{j\to\infty} F\left(l \int_{J_i} x\, dx, l \int_{J_i} \mathbf{u}_j(x)\, dx, l \int_{J_i} \mathbf{u}_j'(x)\, dx\right) \\
&= F\left(l \int_{J_i} x\, dx, l \int_{J_i} \mathbf{u}(x)\, dx, l \int_{J_i} \mathbf{u}'(x)\, dx\right),
\end{aligned}$$

for each i. But for the probability measure $\mu_{i,j}$ determined through the formula

$$\langle \phi, \mu_{i,j} \rangle = l \int_{J_i} \phi(\mathbf{u}_j'(x))\, dx,$$

Jensen's inequality applied to

$$\phi = F\left(l \int_{J_i} x\, dx, l \int_{J_i} \mathbf{u}_j(x)\, dx, \cdot\right)$$

implies

$$F\left(l\int_{J_i} x\,dx, l\int_{J_i} \mathbf{u}_j(x)\,dx, l\int_{J_i} \mathbf{u}_j'(x)\,dx\right) \le$$
$$l\int_{J_i} F\left(l\int_{J_i} x\,dx, l\int_{J_i} \mathbf{u}_j(x)\,dx, \mathbf{u}_j'(x)\right)\,dx.$$

The right-hand side, when m is large, is essentially the integral

$$l\int_{J_i} F\left(x, \mathbf{u}_j(x), \mathbf{u}_j'(x)\right)\,dx.$$

If we put all of our ingredients together, we arrive at the claimed weak lower semicontinuity result

$$\int_J F(x, \mathbf{u}(x), \mathbf{u}'(x))\,dx \le \lim_{j\to\infty}\int_J F(x, \mathbf{u}_j(x), \mathbf{u}_j'(x))\,dx,$$

as the limit for l large, of

$$\begin{aligned}
&\sum_{i=1}^{l}\frac{1}{l}F\left(l\int_{J_i} x\,dx, l\int_{J_i}\mathbf{u}(x)\,dx, l\int_{J_i}\mathbf{u}'(x)\,dx\right)\\
&=\lim_{j\to\infty}\sum_{i=1}^{l}\frac{1}{l}F\left(l\int_{J_i} x\,dx, l\int_{J_i}\mathbf{u}_j(x)\,dx, l\int_{J_i}\mathbf{u}_j'(x)\,dx\right),\\
&\le\lim_{j\to\infty}\sum_{i=1}^{l}\int_{J_i}F\left(l\int_{J_i} x\,dx, l\int_{J_i}\mathbf{u}_j(x)\,dx, \mathbf{u}_j'(x)\right)\,dx\\
&\sim\lim_{j\to\infty}\sum_{i=1}^{l}\int_{J_i}F\left(x, \mathbf{u}_j(x), \mathbf{u}_j'(x)\right)\,dx\\
&=\lim_{j\to\infty}\int_J F\left(x, \mathbf{u}_j(x), \mathbf{u}_j'(x)\right)\,dx.
\end{aligned}$$

4.4 An Existence Result

Based on our main weak lower semicontinuity result Theorem 4.2, we can now prove a principal result for existence of optimal solutions for variational problems

with integral functionals of the kind

$$\text{Minimize in } \mathbf{u} \in \mathcal{A}: \quad I(\mathbf{u}) = \int_J F(x, \mathbf{u}(x), \mathbf{u}'(x))\,dx \tag{4.9}$$

where $\mathcal{A} \subset W^{1,p}(J; \mathbb{R}^n)$ is a non-empty subset encoding constraints to be respected. The exponent p is assumed to be strictly greater than unity, and finite.

Theorem 4.3 *Let*

$$F(x, \mathbf{u}, \mathbf{U}) : J \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

be continuous in $(\mathbf{u}, \mathbf{U})$ *for a.e.* $x \in J$, *and measurable in* x *for every pair* $(\mathbf{u}, \mathbf{U})$. *Suppose that the following three main conditions hold:*

1. *coercivity and growth: there are two positive constant* $C_- \le C_+$, *and a exponent* $p > 1$ *with*

$$C_- \left(|\mathbf{U}|^p - 1\right) \le F(x, \mathbf{u}, \mathbf{U}) \le C_+ \left(|\mathbf{U}|^p + 1\right),$$

for every $(x, \mathbf{u}, \mathbf{U}) \in J \times \mathbb{R}^n \times \mathbb{R}^n$;

2. *convexity: for a.e.* $x \in J$, *and every* $\mathbf{u} \in \mathbb{R}^n$, *we have that*

$$F(x, \mathbf{u}, \cdot) : \mathbb{R}^n \to \mathbb{R}$$

is convex;

3. *the set* $\mathcal{A}$ *is weakly closed, and for some* $x_0 \in J$, *the set of numbers*

$$\{\mathbf{u}(x_0) : \mathbf{u} \in \mathcal{A}\}$$

is bounded.

Then there are optimal solutions for the corresponding variational principle (4.9).

Proof At this stage the proof does not show any surprise, and follows the guide of the direct method Proposition 3.1. If m is the value of the associated infimum, the coercivity and growth condition implies that $m \in \mathbb{R}$. Let $\{\mathbf{u}_j\}$ be a minimizing sequence. Through the coercivity condition, we can ensure that, possibly for a suitable subsequence, the sequence of derivatives $\{\mathbf{u}'_j\}$ is bounded in $L^p(J; \mathbb{R}^n)$. Indeed, for such minimizing sequence we would have, thanks to the coercivity condition,

$$\|u'_j\|^p_{L^p(J;\mathbb{R}^N)} \le 1 + \frac{1}{C_-} I(u_j).$$

Since $I(u_j) \searrow m \in \mathbb{R}$, this last inequality implies that the sequence of derivatives $\{u'_j\}$ is uniformly bounded in $L^p(J; \mathbb{R}^N)$. If we write

$$\mathbf{u}_j(x) = \mathbf{u}_j(x_0) + \int_{x_0}^{x} \mathbf{u}'_j(s)\, ds,$$

the condition assumed on the point $x_0 \in J$ and all feasible paths in $\mathcal{A}$, together with the previous uniform bound on derivatives, imply that $\{\mathbf{u}_j\}$ is uniformly bounded in $W^{1,p}(J; \mathbb{R}^n)$, and hence $\mathbf{u}_j \rightharpoonup \mathbf{u}$ for some $\mathbf{u}$, which belongs to $\mathcal{A}$, if this subset is weakly closed. Our weak lower semicontinuity result Theorem 4.2 guarantees that

$$m \le I(\mathbf{u}) \le \liminf_{j \to \infty} I(\mathbf{u}_j) \le m,$$

and $\mathbf{u}$ becomes a true minimizer for our problem. □

There are several remarks worth considering.

1. We know that the convexity condition on the variable $\mathbf{U}$ is unavoidable for weak lower semicontinuity, however both the coercivity condition and the weak closeness of the feasible set $\mathcal{A}$ may come in various forms. In particular, only the coercivity inequality

$$C\left(|\mathbf{U}|^p - 1\right) \le F(x, \mathbf{u}, \mathbf{U}) \tag{4.10}$$

is necessary for an existence result. However, since in such a situation the possibility that I may take on the value $+\infty$ somewhere in the space $W^{1,p}(J; \mathbb{R}^n)$ is possible, one needs to make sure that the feasible set $\mathcal{A}$, where I is supposed to take on finite values, is non-empty, i.e. I is not identically $+\infty$, and then

$$m = \inf_{\mathbf{u} \in \mathcal{A}} I(\mathbf{u}),$$

is a real number. In this case, we can talk about minimizing sequences $\{\mathbf{u}_j\}$ with $I(\mathbf{u}_j) \searrow m \in \mathbb{R}$, and proceed with the same proof of Theorem 4.3.

2. The coercivity condition in Theorem 4.3 is too rigid in practice. It can be replaced in the statement of that result by the more flexible one that follows, and the conclusion is valid in the same way: there is $C > 0$ and a exponent $p > 1$ such that

$$C\left(\|\mathbf{u}'\|_{L^p(J;\mathbb{R}^N)} - 1\right) \le I(\mathbf{u}). \tag{4.11}$$

3. The uniqueness of minimizers can only be achieved under much more restrictive conditions as it is not sufficient to strengthen the convexity of the integrand $F(x, \mathbf{u}, \mathbf{U})$ with respect to $\mathbf{U}$ to strict convexity. This was already indicated in the comments around Proposition 3.3. In fact, the following statement is a direct consequence of that result.

Proposition 4.2 *In addition to all assumptions in Theorem 4.3, suppose that the integrand* $F(x, \mathbf{u}, \mathbf{U})$ *is jointly strictly convex with respect to pairs* $(\mathbf{u}, \mathbf{U})$. *Then there is a unique minimizer for problem* (4.9).

However, just as we have noticed above concerning coercivity, some times the direct application of Proposition 3.3 may be more flexible than Proposition 4.2.

4.5 Some Examples

Convexity and coercivity are the two main ingredients that guarantee the existence of optimal solutions for standard variational problems according to Theorem 4.3. We will proceed to cover some examples without further comment on these two fundamental properties of functions. Later, when dealing with higher-dimensional problems that are often more involved, we will try to be more systematic.

We examine next several examples with care. In many cases, the convexity requirement is easier to check than the coercivity.

Example 4.1 An elastic string of unit length is supported on its two end-points at the same height $y = 0$. Its equilibrium profile under the action of a vertical load of density $g(x)$ for $x \in [0, 1]$ is the one minimizing internal energy which is approximated, given that the values of y' are assumed to be reasonably small, by the functional

$$E(y) = \int_0^1 [\frac{1}{2}y'(x)^2 + g(x)y(x)]\,dx.$$

We assume $g \in L^2(0, 1)$. For the end-point conditions

$$y(0) = y(1) = 0,$$

existence of minimizers can be achieved through our main existence theorem, though the coercivity hypothesis requires some work. In fact, the initial form of coercivity in Theorem 4.3 can hardly be used directly. However, it is possible to establish (4.11) for $p = 2$. As a matter of fact, the ideas that follow can be used in many situations to properly adjust the coercivity condition. From

$$y(x) = \int_0^x y'(s)\,ds$$

it is easy to find, through Hölder's inequality, that

$$y(x)^2 \leq x \int_0^1 y'(s)^2\,ds,$$

and

$$\int_0^1 y(x)^2\,dx \le \int_0^1 y'(x)^2\,dx.$$

It is even true that

$$\int_0^1 y(x)^2\,dx \le \frac{1}{2}\int_0^1 y'(x)^2\,dx,$$

but this better estimate does not mean any real improvement on the coercivity condition we are seeking. Then for the second term in $E(y)$, we can write, for arbitrary $\epsilon > 0$,

$$\begin{aligned}\int_0^1 |g(x)|\,|y(x)|\,dx \le &\frac{1}{2\epsilon^2}\int_0^1 g(x)^2\,dx + \frac{\epsilon^2}{2}\int_0^1 y(x)^2\,dx\\ =&\frac{1}{2\epsilon^2}\int_0^1 g(x)^2\,dx + \frac{\epsilon^2}{2}\int_0^1 y'(x)^2\,dx.\end{aligned}$$

We are using the inequality

$$ab = (\epsilon a)\left(\frac{b}{\epsilon}\right) \le \frac{1}{2}\epsilon^2 a^2 + \frac{1}{2\epsilon^2}b^2, \tag{4.12}$$

valid for arbitrary positive numbers ϵ, a, b. If we take this estimate to the functional $E(y)$, we see that

$$\begin{aligned}E(y) \ge&\frac{1}{2}\int_0^1 y'(x)^2\,dx - \int_0^1 |g(x)|\,|y(x)|\,dx\\ \ge&\frac{1-\epsilon^2}{2}\int_0^1 y'(x)^2\,dx - \frac{1}{2\epsilon^2}\int_0^1 g(x)^2\,dx\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we can take, say, $2\epsilon^2 = 1$, and then

$$E(y) \ge \frac{1}{4}\int_0^1 y'(x)^2\,dx - \int_0^1 g(x)^2\,dx.$$

This inequality implies the necessary coercivity. Though the integrand

$$F(x, y, Y) = \frac{1}{2}Y^2 + g(x)y$$

is not jointly strictly convex with respect to (y, Y), and hence we cannot apply directly Proposition 4.2, it is not difficult to argue, through Proposition 3.3, that

there is a unique minimizer because the functional, not the integrand,

$$E : H_0^1(0,1) \to \mathbb{R}$$

is strictly convex.

Example 4.2 A much-studied scalar uni-dimensional problem is of the form

$$I(u) = \int_0^1 \left[a u'(x)^2 + g(u(x)) \right] dx, \quad a > 0, \tag{4.13}$$

under typical end-point conditions, with a real function g which is assumed to be just continuous. For simplicity, we also take vanishing end-point conditions. Once again coercivity becomes the main point to be clarified for the application of Theorem 4.3. Uniqueness is, however, compromised unless convexity conditions are imposed on the function g. As in the preceding example, we see that

$$|u(x)| \le \sqrt{x} \left(\int_0^1 u'(s)^2 \, ds \right)^{1/2} \le \|u'\|_{L^2(0,1)}.$$

The second term in $I(u)$ is bounded by

$$\max_{U \in J_u} |g(U)|, \quad J_u = [-\|u'\|_{L^2(0,1)}, \|u'\|_{L^2(0,1)}].$$

Suppose that

$$|g(U)| \le b|U| + C, \quad b < a, C \in \mathbb{R}, U \in J_u.$$

Then the previous maximum can be bounded by

$$b\|u'\|_{L^2(0,1)} + C,$$

and then

$$I(u) \ge (a - b)\|u'\|_{L^2(0,1)} - C.$$

Again, this inequality fits better (4.11) than the coercivity condition in Theorem 4.3. This last theorem ensures that I in (4.13) admits global minimizers.

Example 4.3 In the context of Sect. 1.3.6, the integrand

$$L(\mathbf{x}, \mathbf{x}') = \frac{1}{2} m |\mathbf{x}'|^2 - \frac{1}{2} k |\mathbf{x}|^2, \quad m, k > 0,$$

is the Lagrangian for a simple harmonic oscillator. The corresponding integral is the action integral. In these problems in Mechanics only the initial conditions are enforced so that feasible paths are required to comply with

$$\mathbf{x}(0) = \mathbf{x}_0, \quad \mathbf{x}'(0) = \mathbf{x}'_0,$$

for given vectors $\mathbf{x}_0, \mathbf{x}'_0 \in \mathbb{R}^N$. We consider then the variational problem

$$\text{Minimize in } \mathbf{x} \in \mathcal{A}: \quad A(\mathbf{x}) = \int_0^T L(\mathbf{x}(t), \mathbf{x}'(t))\, dt \tag{4.14}$$

where this time

$$\mathcal{A} = \left\{ \mathbf{x} \in H^2(0, T; \mathbb{R}^N) : \mathbf{x}(0) = \mathbf{x}_0, \int_0^\epsilon |\mathbf{x}'(t) - \mathbf{x}'_0|^2\, dt \to 0 \text{ as } \epsilon \to 0 \right\}.$$

Without loss of generality, for the sake of simplicity, we can take

$$\mathbf{x}_0 = \mathbf{x}'_0 = \mathbf{0},$$

and in this case $\mathcal{A}$ becomes a vector space.

We would like to show that the value of the minimum m in (4.14) vanishes, and that the only minimizer is the trivial path $\mathbf{x} \equiv \mathbf{0}$. This requires some computations to assess the relative size of the two terms in the Lagrangian. Using ideas as in the previous examples, we can write

$$|\mathbf{x}(t)|^2 \le t \int_0^t |\mathbf{x}'(s)|^2\, ds.$$

From this inequality, and by using Fubini's theorem,

$$\begin{aligned}
\int_0^T |\mathbf{x}(t)|^2\, dt &\le \int_0^T t \int_0^t |\mathbf{x}'(s)|^2\, ds\, dt \\
&= \int_0^T \int_s^T t \int_0^t |\mathbf{x}'(s)|^2\, dt\, ds \\
&= \int_0^T \frac{T^2 - s^2}{2} |\mathbf{x}'(s)|^2\, ds.
\end{aligned}$$

If, relying on this estimate, we compare the two terms in the action integral $A(\mathbf{x})$, we would find

$$\begin{aligned}\int_0^T \left(\frac{m}{k}|\mathbf{x}'(t)|^2 - |\mathbf{x}(t)|^2\right) dt &\geq \int_0^T \left(\frac{m}{k} - \frac{T^2 - t^2}{2}\right) |\mathbf{x}'(t)|^2 \, dt \\ &\geq \left(\frac{m}{k} - \frac{T^2}{2}\right) \int_0^T |\mathbf{x}'(t)|^2 \, dt.\end{aligned}$$

Since the action integral has to be minimized almost instantaneously as the movement progresses, we can assume that the time horizon T is small, namely, if

$$T < \sqrt{\frac{2m}{k}}, \tag{4.15}$$

then $A(\mathbf{x})$ is non-negative on $\mathcal{A}$, and it can only vanish if $\mathbf{x}' \equiv \mathbf{0}$, i.e. $\mathbf{x} \equiv \mathbf{0}$ in $[0, T]$. For larger values of T, one can proceed in successive subintervals of length smaller than the right-hand side in (4.15). This shows that once the pendulum touches the rest conditions, it needs an external agent to start again the movement.

Example 4.4 The behavior of non-linear, elastic materials (rubber for instance) is much more complicated to understand and simulate than their linear counterparts. Suppose that an elastic stick is of length one in its natural rest configuration. When subjected to an elongation of its right-end point $x = 1$ to some other value, say, $L > 0$, maintaining $x = 0$ in its position, it will respond seeking its minimum internal energy under the new end-point conditions. We postulate that each possible deformation of the bar is described by a function

$$u(x) : (0, 1) \to \mathbb{R}, \quad u(0) = 0, u(1) = L, \quad u'(x) > 0,$$

with internal potential energy given by the functional

$$E(u) = \int_0^1 \left[\frac{1}{2}u'(x)^2 + \frac{\alpha}{u'(x)} + g(x)u(x)\right] dx, \quad \alpha > 0,$$

where the second term aims at assigning an infinite energy required to compress some part of the bar to zero volume, and to avoid interpenetration of matter $u' \leq 0$; and the third one accounts for a bulk load with density $g(x)$. The optimal solutions of the variational problem

$$\text{Minimize in } u(x) \in H^1(0, 1) : \quad E(u)$$

under

$$u(0) = 0, \quad u(1) = L,$$

will be taken to be as configurations of minimal energy, and they will be the ones that can be adopted by the bar under the indicated end-point conditions. It is elementary to check that the integrand

$$F(x, u, U) = \frac{1}{2}U^2 + \frac{\alpha}{U} + g(x)u$$

is strictly convex with respect to U in the positive part $U > 0$, linear with respect to u, while the functional $E(u)$ is coercive in $H^1(0, 1)$, just as one of the examples above. We can therefore conclude that there is exactly one minimizer of the problem.

Unfortunately, many of the geometrical or physical variational problems, some of which were introduced in the initial chapter, do not fall under the action of our main existence theorem because the exponent p for the coercivity in that statement is assumed greater than 1, and quite often functionals with a geometric meaning have only linear growth on $\mathbf{u}'$. The trouble with $p = 1$ in our main result Theorem 4.3 is intimately related to the difficulty explained in Example 2.9. The coercivity condition in the statement of that theorem would lead to a uniform bound of a minimizing sequence in $W^{1,1}(J; \mathbb{R}^N)$, but this is not sufficient to ensure a limit function which would be a candidate for minimizer. In such cases, an additional argument of some kind ought to be invoked for the existence of a minimizer; or else, one can examine the indirect method (ideas in Sect. 3.7).

Example 4.5 Let us deal with the problem for the brachistochrone described in Sect. 1.3. The problem is

$$\text{Minimize in } u(x): \quad B(u) = \int_0^A \frac{\sqrt{1 + u'(x)^2}}{\sqrt{x}}\, dx$$

subject to

$$u(0) = 0, \quad u(A) = a.$$

The factor $1/\sqrt{x}$, even though blows up at $x = 0$, is not a problem as we would have a coercivity constant of tremendous force. The difficulty lies with the linear growth on u' of the integrand. For some minimizing sequence $\{u_j\}$, we need to rule out the possibility shown in Example 2.9. Unfortunately, there is no general rule for this since each particular example may require a different idea, if indeed there are minimizing sequences not concentrating. This would be impossible for a variational problem for which concentration of derivatives is “essential” to the minimization process. Quite often, the geometrical or physical meaning of the situation may lead to a clear argument about the existence of an optimal solution. In this particular instance, we can interpret the problem as a shortest distance problem with respect

to the measure

$$dm = \frac{dx}{\sqrt{x}}.$$

Note that this measure assigns distances in the real line in the form

$$\int_{\delta}^{\epsilon} \frac{dx}{\sqrt{x}} = 2(\sqrt{\epsilon} - \sqrt{\delta}).$$

Therefore our problem consists in finding the shortest path, the geodesic with respect to this new way of measuring distances, and minimal distance between the points $(0, 0)$ and (A, a). As such, there should be an optimal solution. We will later calculate it.

Example 4.6 Suppose

$$\mathbf{F}(\mathbf{x}) : \mathbb{R}^3 \to \mathbb{R}^3$$

is a smooth vector field in three-dimensional space, representing a certain force field. The work done by it when a unit-mass particle goes from a point $\mathbf{x}(0) = \mathbf{x}_0$ to $\mathbf{x}(1) = \mathbf{x}_1$ is measured by the path integral

$$W(\mathbf{x}) = \int_0^1 \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\, dt.$$

We well know that if $\mathbf{F}$ is conservative so that there is an underlying potential

$$\phi(\mathbf{x}) : \mathbb{R}^3 \to \mathbb{R}, \quad \mathbf{F} = \nabla\phi,$$

then such work is exactly the potential difference

$$\begin{aligned}\int_0^1 \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\, dt =& \int_0^1 \frac{d}{dt}[\phi(\mathbf{x}(t))]\, dt \\ =& \phi(\mathbf{x}_1) - \phi(\mathbf{x}_0),\end{aligned}$$

and so the work functional $W(\mathbf{x})$ is independent of the path $\mathbf{x}$. In the jargon of variational problems, we would say that W is a null-lagrangian. But if $\mathbf{F}$ is non-conservative, then $W(\mathbf{x})$ will depend on $\mathbf{x}$, and hence one may wonder, in principle, about the cheapest path in terms of work done through it. We would like, hence, to consider the variational problem

$$\text{Minimize in } \mathbf{x}(t) : \quad W(\mathbf{x}) = \int_0^1 \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\, dt$$

under the given end-point conditions. The integrand corresponding to the problem is

$$F(\mathbf{x}, \mathbf{X}) = \mathbf{F}(\mathbf{x}) \cdot \mathbf{X}$$

which is linear in the derivative variable $\mathbf{X}$. The natural functional space for the problem is then $W^{1,1}((0, 1); \mathbb{R}^N)$. Paths $\mathbf{x}(t)$ in this space are absolutely continuous and $W(\mathbf{x})$ is well-defined for them. Can we ensure that there is always paths with minimal work joining any two points $\mathbf{x}_0$, $\mathbf{x}_1$? Theorem 4.3 cannot be applied, and there does not seem to be any valid argument available to ensure the existence of optimal paths. We will look at the problem below from the point of view of optimality (indirect method).

4.5.1 Existence Under Constraints

In a good proportion of situations, one is interested in minimizing a certain integral functional under additional constraints other than fixed end-point conditions. These can be of a global nature, quite often expressed through an additional integral functional; or of a local nature by demanding that constraints are respected in a pointwise fashion. At the level of existence of optimal solutions such constraints can always be incorporated into the feasible set $\mathcal{A}$ of our main result Theorem 4.3. We present two examples to cover those two situations.

Example 4.7 In the context of Example 4.1 above, where the profile adopted by a linear, elastic string subjected to a vertical load is given by the minimizer of the functional

$$E(y) = \int_0^1 [\frac{1}{2} y'(x)^2 + g(x)y(x)]\, dx$$

under end-point constraints, we may complete the situation with an additional interesting ingredient. Suppose that there is a rigid obstacle whose profile is represented by the continuous function

$$\phi(x) : (0, 1) \to \mathbb{R}$$

in such a way that feasible configurations $y(x)$ must comply with

$$y(x) \geq \phi(x), \quad x \in (0, 1).$$

In this case, the admissible set for the optimization problem is

$$\mathcal{A} = \{y(x) \in H^1(0, 1) : y(x) \geq \phi(x), y(0) = y_0, y(1) = y_1\},$$

and obviously we need to enforce

$$\phi(0) < y_0, \quad \phi(1) < y_1,$$

if we want $\mathcal{A}$ to be non-empty. Theorem 4.3 can be applied. The only new feature, compared to Example 4.1, is to check if the feasible set of this new situation is weakly closed. But since weak convergence in $H^1(0,1)$ implies convergence in $L^\infty(0,1)$, an inequality not involving derivatives like

$$y_j(x) \geq \phi(x)$$

is preserved if $y_j \to y$ uniformly in $(0,1)$. Such problems have been studied quite a lot in the literature and are referred to as obstacle problems. (See Exercise 29).

Example 4.8 We recover here the problem of the hanging cable introduced in the first chapter. We are concerned with minimizing the functional

$$\int_0^H u(x)\sqrt{1+u'(x)^2}\,dx$$

under the conditions

$$u(0) = u(H) = 0, \quad L = \int_0^H \sqrt{1+u'(x)^2}\,dx,$$

and $0 < H < L$. Since both integrands involved in the two functionals share linear growth in the derivative variable, the natural space in which to setup the problem is $W^{1,1}(0,H)$. In addition, the global integral constraint must be part of the feasible set of competing functions

$$\mathcal{A} = \left\{u(x) \in W^{1,1}(0,H) : u(0) = u(H) = 0, L = \int_0^H \sqrt{1+u'(x)^2}\,dx\right\}.$$

There are two main difficulties for the use of Theorem 4.3 in this situation: one is the coercivity in which the growth exponent $p = 1$ for integrands keeps us from using it directly; the other one is the weakly closeness of $\mathcal{A}$ which is not correct either. Can readers provide an explicit sequence $\{u_j\} \subset \mathcal{A}$ (for $H = 1$ and $L = 2/\sqrt{2}$) such that its weak limit u does not belong to $\mathcal{A}$? (Exercise 14). Even so, the physical interpretation of the problem clearly supports the existence of a unique optimal solution which, after all, would be the profile adopted by the cable in practice.

4.6 Optimality Conditions

In abstract terms, optimality conditions require to work in a Hilbert space to be able to compute derivatives as in Lemma 2.7. This would lead us to restrict attention to the particular case in which the exponent for coercivity is $p = 2$ in Theorem 4.3, so that the corresponding integral functional would admit minimizing sequences in the Hilbert space $H^1(J; \mathbb{R}^N)$. This is another aspect in which studying integral functional becomes much more flexible than doing so for classes of functionals not specifying its nature. One can actually look at optimality conditions in a general Sobolev space $W^{1,p}(J; \mathbb{R}^N)$ for any value of $p > 1$, under appropriate assumptions. Yet, since, typically, optimality demands more restrictive working conditions and many more technicalities, we will be contented with a general discussion in the Hilbert space scenario for growth exponent $p = 2$, and try to expand the scope of this particular situation to other cases through examples. Under a simple coercivity condition like (4.11), where one could allow functionals whose finite set, the set of functions where it attains a finite value, is not a vector-space, optimality is specially delicate.

We will therefore be contented with examining optimality under the action of Theorem 4.3 for the particular case $p = 2$. Namely, the following is a corollary of that theorem with the only difference in the statement that we restrict attention to the case $p = 2$ so that the underlying functional space is the Hilbert space $H^1(J; \mathbb{R}^n)$.

Corollary 4.1 *Let*

$$F(x, \mathbf{u}, \mathbf{U}) : J \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \tag{4.16}$$

be continuous in $(\mathbf{u}, \mathbf{U})$ *for a.e.* $x \in J$*, and measurable in* x *for every pair* $(\mathbf{u}, \mathbf{U})$*. Suppose that the following three main conditions hold:*

1. *coercivity and growth: there are positive constants* $C_- \le C_+$ *with*

$$C_- \left(|\mathbf{U}|^2 - 1\right) \le F(x, \mathbf{u}, \mathbf{U}) \le C_+ \left(|\mathbf{U}|^2 + 1\right),$$

   for $(x, \mathbf{u}, \mathbf{U}) \in J \times \mathbb{R}^n \times \mathbb{R}^n$*;*
2. *convexity: for a.e.* $x \in J$*, and every* $\mathbf{u} \in \mathbb{R}^n$*, we have that*

$$F(x, \mathbf{u}, \cdot) : \mathbb{R}^n \to \mathbb{R}$$

   is convex;
3. *the set*

$$\mathcal{A} \subset H^1(J; \mathbb{R}^n)$$

is weakly closed, and for some $x_0 \in J$*, the set of numbers*

$$\{\mathbf{u}(x_0) : \mathbf{u} \in \mathcal{A}\}$$

is bounded.

Then there are optimal solutions for the variational problem

$$\textit{Minimize in } \mathbf{u} \in \mathcal{A} : \quad I(\mathbf{u}) = \int_J F(x, \mathbf{u}(x), \mathbf{u}'(x))\, dx. \tag{4.17}$$

We place the following discussion in the context of Sect. 2.11.4, and assume that the integrand $F(x, \mathbf{u}, \mathbf{U})$ in (4.16) is, in addition to hypotheses in the previous corollary, smooth (at least C^1) with respect to pairs

$$(\mathbf{u}, \mathbf{U}) \in \mathbb{R}^n \times \mathbb{R}^n$$

for a.e. $x \in J$. Moreover, as we describe the process, we will be putting aside assumptions as they are needed, without pretending to isolate the most general setting.

One of the most studied situations is concerned with end-point boundary value problems in which, in the context of Theorem 4.3 or Corollary 4.1,

$$\mathcal{A} = \{\mathbf{u} \in H^1(J; \mathbb{R}^n) : \mathbf{u}(x_0) = \mathbf{u}_0, \mathbf{u}(x_1) = \mathbf{u}_1\} \tag{4.18}$$

for two fixed vectors $\mathbf{u}_0, \mathbf{u}_1 \in \mathbb{R}^n$. Note how the third main condition in Theorem 4.3 or Corollary 4.1 is automatically fulfilled. Suppose we have found, through Corollary 4.1 or otherwise, a minimizer $\mathbf{u} \in \mathcal{A}$. The basic idea was stated in Proposition 3.4. We would like to "make variations" based on $\mathbf{u}$ by perturbing it with $\mathbf{v}$ in the underlying subspace

$$H_0^1(J; \mathbb{R}^n) = \{\mathbf{v} \in H^1(J; \mathbb{R}^n) : \mathbf{v}(x_0) = \mathbf{v}(x_1) = \mathbf{0}\}$$

that has been introduced earlier in the text. The combination $\mathbf{u} + s\mathbf{v}$ for arbitrary real s is a feasible path for our variational problem, and then we should have

$$I(\mathbf{u}) \leq I(\mathbf{u} + s\mathbf{v})$$

for all real s. This means that the real function

$$s \mapsto I(\mathbf{u} + s\mathbf{v}) \tag{4.19}$$

has a global minimum for $s = 0$, and hence

$$\left.\frac{dI(\mathbf{u} + s\mathbf{v})}{ds}\right|_{s=0} = 0,$$

provided (4.19) is differentiable. We can differentiate formally under the integral sign to find

$$\int_J \left[F_{\mathbf{u}}(x, \mathbf{u}(x), \mathbf{u}'(x)) \cdot \mathbf{v}(x) + F_{\mathbf{U}}(x, \mathbf{u}(x), \mathbf{u}'(x)) \cdot \mathbf{v}'(x)\right] dx = 0 \tag{4.20}$$

for every such $\mathbf{v} \in H_0^1(J; \mathbb{R}^n)$. If we recall that functions in $H^1(J; \mathbb{R}^n)$ are continuous and uniformly bounded, for the integrals in (4.20) to be well-defined and so (4.19) be differentiable, we only require, for some positive constant C, that

$$|F_{\mathbf{u}}(x, \mathbf{u}, \mathbf{U})| \le C(|\mathbf{U}|^2 + 1), \quad |F_{\mathbf{U}}(x, \mathbf{u}, \mathbf{U})| \le C(|\mathbf{U}| + 1).$$

Consider the orthogonal complement, in the Hilbert space $H^1(J; \mathbb{R}^n)$, to the subspace $\mathbf{e}$ generated by the constant vectors $\mathbf{e}_i$ of the canonical basis of $\mathbb{R}^n$, i.e.

$$\mathbf{e}^{\perp} = \{\mathbf{w}(x) \in H^1(J; \mathbb{R}^n) : \int_J \mathbf{w}(x)\, dx = \mathbf{0}\}.$$

For every $\mathbf{w} \in \mathbf{e}^{\perp}$, its primitive

$$\mathbf{v}(x) = \int_{x_0}^{x} \mathbf{w}(y)\, dy$$

belongs to $H_0^1(J; \mathbb{R}^n)$ and

$$\mathbf{v}'(x) = \mathbf{w}(x), \quad x \in J.$$

On the other hand, if we let

$$\Phi(x) = \int_{x_0}^{x} F_{\mathbf{u}}(y, \mathbf{u}(y), \mathbf{u}'(y))\, dy, \quad \Psi(x) = F_{\mathbf{U}}(x, \mathbf{u}(x), \mathbf{u}'(x)),$$

(4.20) can be recast in the form

$$\int_J [\Phi'(x) \cdot \mathbf{v}(x) + \Psi(x) \cdot \mathbf{v}'(x)]\, dx = 0.$$

Bearing in mind that

$$\mathbf{v}(x_0) = \mathbf{v}(x_1) = \mathbf{0},$$

an integration by parts in the first term leads to

$$\int_J [\Phi(x) - \Psi(x)] \cdot \mathbf{w}(x)\, dx = 0$$

for every $\mathbf{w} \in \mathbf{e}^\perp$. Conclude that

$$\Phi(x) + \Psi(x) = \mathbf{c}, \text{ a constant vector in } J,$$

or by differentiation

$$F_\mathbf{u}(x, \mathbf{u}(x), \mathbf{u}'(x)) - \frac{d}{dx} F_\mathbf{U}(x, \mathbf{u}(x), \mathbf{u}'(x)) = \mathbf{0} \text{ in } J.$$

We have just proved the following important statement.

Theorem 4.4 *Suppose that the integrand $F(x, \mathbf{u}, \mathbf{U})$ is C^1 with respect to pairs $(\mathbf{u}, \mathbf{U}) \in \mathbb{R}^n \times \mathbb{R}^n$, and that*

$$C_- \left(|\mathbf{U}|^2 - 1\right) \le F(x, \mathbf{u}, \mathbf{U}) \le C_+ \left(|\mathbf{U}|^2 + 1\right),$$

$$|F_\mathbf{u}(x, \mathbf{u}, \mathbf{U})| \le C_+(|\mathbf{U}|^2 + 1), \quad |F_\mathbf{U}(x, \mathbf{u}, \mathbf{U})| \le C_+(|\mathbf{U}| + 1),$$

for some positive constants $C_- \le C_+$, and every $(x, \mathbf{u}, \mathbf{U}) \in J \times \mathbb{R}^n \times \mathbb{R}^n$. Let $\mathbf{u} \in \mathcal{A}$, given in (4.18), be a minimizer of our variational problem (4.17). Then

$$F_\mathbf{U}(x, \mathbf{u}(x), \mathbf{u}'(x))$$

is absolutely continuous in J, and

$$F_\mathbf{u}(x, \mathbf{u}(x), \mathbf{u}'(x)) - \frac{d}{dx} F_\mathbf{U}(x, \mathbf{u}(x), \mathbf{u}'(x)) = \mathbf{0} \text{ a.e. } x \text{ in } J. \tag{4.21}$$

The differential system in this statement is universally known as the Euler-Lagrange system of optimality. In the scalar case $n = 1$, the second-order differential equation in this result together with the end-point conditions is referred to as a Sturm-Liouville problem.

One main application of this result stresses the fact that if the minimizer $\mathbf{u}$ is the outcome of Corollary 4.1, because the suitable hypotheses hold, then $\mathbf{u}$ must be a solution of the corresponding Euler-Lagrange problem (4.21). This is the variational method to show existence of solutions of (4.21). However, some times one might be interested in going the other way, i.e. from solutions of (4.21), assuming that this is possible, to minimizers for (4.17). The passage from minimizers $\mathbf{u}$ for (4.17) to solutions of (4.21) does not require convexity, and the process is known as the necessity of optimality conditions expressed in (4.21) for minimizers; for the passage from solutions $\mathbf{u}$ of (4.21) to minimizers of (4.17) convexity is unavoidable,

and the process is labeled as sufficiency of the Euler-Lagrange problem (4.21) for minimizers of the variational problem (4.17). Though this discussion is formally like the one in Proposition 3.4, in the case of integral functionals the result can be much more explicit, to the point that we can work in the most general Sobolev space $W^{1,1}(J;\mathbb{R}^n)$.

Theorem 4.5 *Let the integrand $F(x, \mathbf{u}, \mathbf{U})$, as above, be C^1- and convex in pairs $(\mathbf{u}, \mathbf{U})$. Suppose the path*

$$\overline{\mathbf{u}}(x) \in W^{1,1}(J;\mathbb{R}^n)$$

is such that (4.21) holds

$$F_{\mathbf{u}}(x, \overline{\mathbf{u}}(x), \overline{\mathbf{u}}'(x)) - \frac{d}{dx} F_{\mathbf{U}}(x, \overline{\mathbf{u}}(x), \overline{\mathbf{u}}'(x)) = \mathbf{0} \text{ a.e. } x \text{ in } J, \tag{4.22}$$

the two terms in this system belong to $L^1(J;\mathbb{R}^n)$, and

$$\overline{\mathbf{u}}(x_0) = \mathbf{u}_0, \quad \overline{\mathbf{u}}(x_1) = \mathbf{u}_1,$$

with $\mathbf{u}_0, \mathbf{u}_1 \in \mathbb{R}^n$. Then $\overline{\mathbf{u}}$ is a minimizer for the problem

$$\text{Minimize in } \mathbf{u} \in \mathcal{A}: \quad I(\mathbf{u}) = \int_J F(x, \mathbf{u}(x), \mathbf{u}'(x))\, dx \tag{4.23}$$

in the feasible set

$$\mathcal{A} = \{\mathbf{u} \in W^{1,1}(J;\mathbb{R}^n) : \mathbf{u}(x_0) = \mathbf{u}_0, \mathbf{u}(x_1) = \mathbf{u}_1\}. \tag{4.24}$$

Proof Let

$$\mathbf{v} \in W_0^{1,1}(J;\mathbb{R}^n),$$

which is the subspace of $W^{1,1}(J;\mathbb{R}^n)$ in (4.24) corresponding to $\mathbf{u}_0 = \mathbf{u}_1 = \mathbf{0}$. We know that $\mathbf{v} \in L^\infty(J;\mathbb{R}^n)$ and, moreover, the formula of integration by parts (2.22) guarantees that

$$\int_J \frac{d}{dx} F_{\mathbf{U}}(x, \overline{\mathbf{u}}(x), \overline{\mathbf{u}}'(x)) \cdot \mathbf{v}(x)\, dx = -\int_J F_{\mathbf{U}}(x, \overline{\mathbf{u}}(x), \overline{\mathbf{u}}'(x)) \cdot \mathbf{v}'(x)\, dx$$

is correct. This identity implies, through (4.22), that

$$\int_J \left[F_{\mathbf{u}}(x, \overline{\mathbf{u}}(x), \overline{\mathbf{u}}'(x)) \cdot \mathbf{v}(x) + F_{\mathbf{U}}(x, \overline{\mathbf{u}}(x), \overline{\mathbf{u}}'(x)) \cdot \mathbf{v}'(x)\right] dx = 0. \tag{4.25}$$

On the other hand, it is well-known that a characterization of convexity for C^1-functions, like $F(x, \cdot, \cdot)$, for a.e. $x \in J$, is

$$F(x, \mathbf{u}+\mathbf{v}, \mathbf{U}+\mathbf{V}) \geq F(x, \mathbf{u}, \mathbf{U}) + F_{\mathbf{u}}(x, \mathbf{u}, \mathbf{U}) \cdot \mathbf{v} + F_{\mathbf{U}}(x, \mathbf{u}, \mathbf{U}) \cdot \mathbf{V}$$

for every $\mathbf{u}, \mathbf{U}, \mathbf{v}, \mathbf{V} \in \mathbb{R}^n$. In particular, for a.e. $x \in J$,

$$\begin{aligned} F(x, \overline{\mathbf{u}}(x) + \mathbf{v}(x), \overline{\mathbf{u}}'(x) + \mathbf{v}'(x)) \geq & F(x, \overline{\mathbf{u}}(x), \overline{\mathbf{u}}'(x)) \\ & + F_{\mathbf{u}}(x, \overline{\mathbf{u}}(x), \overline{\mathbf{u}}'(x)) \cdot \mathbf{v}(x) \\ & + F_{\mathbf{U}}(x, \overline{\mathbf{u}}(x), \overline{\mathbf{u}}'(x)) \cdot \mathbf{v}'(x). \end{aligned}$$

Upon integration on $x \in J$, bearing in mind (4.25), we find

$$I(\overline{\mathbf{u}} + \mathbf{v}) \geq I(\overline{\mathbf{u}}).$$

The arbitrariness of $\mathbf{v} \in W_0^{1,1}(J; \mathbb{R}^n)$ implies our conclusion. □

As usual, uniqueness of minimizer is associated with some form of strict convexity.

Proposition 4.3 *Assume, in addition to hypotheses in Theorem 4.5, that the integrand $F(x, \mathbf{u}, \mathbf{U})$ is jointly convex on pairs $(\mathbf{u}, \mathbf{U})$. Then the path $\overline{\mathbf{u}}$ in Theorem 4.5 is the unique minimizer of the problem* (4.23)–(4.24).

The proof is just like other such results, for example Proposition 4.2.

4.7 Some Explicit Examples

The use of optimality conditions in specific examples, may serve, among other important purposes beyond the scope of this text, to two main objectives: one is to try to find minimizers when the direct method delivers them; another one, is to explore if existence of minimizers can be shown once we have calculated them in the context of the indirect method (Sect. 3.7). We will only consider end-point conditions as additional constraints in the feasible set $\mathcal{A}$; other kind of constraints ask for the introduction of multipliers, and are a bit beyond the scope of this text. See some additional comments below in Example 4.11.

Example 4.9 We start by looking at the first examples of Sect. 4.5. It is very easy to conclude that the unique minimizer $y(x)$ for Example 4.1 is the unique solution of the differential problem

$$-y''(x) + g(x) = 0 \text{ in } (0, 1), \quad y(0) = y(1) = 0. \tag{4.26}$$

It is elementary to find that such a solution is given explicitly by the formula

$$y(x) = \int_0^x (x-s)g(s)\,ds - x\int_0^1 (1-s)g(s)\,ds. \tag{4.27}$$

For Example 4.2, we find the optimality differential equation

$$-2au''(x) + g'(u(x)) = 0 \text{ in } (0,1), \quad u(0) = u(1) = 0.$$

This time is impossible to know, without looking at the functional it is coming from, if this equation admits solutions. Note that the problem is non-linear as soon as g is non-quadratic. According to our discussion above, we can certify that this non-linear differential problem admits solutions (possibly in a non-unique way) whenever

$$|g(u)| \le b|u| + C, \quad b < a, C \in \mathbb{R}.$$

For Example 4.3, there is no much to say. Optimality conditions become a differential system

$$m\mathbf{x}''(t) + k\mathbf{x}(t) = \mathbf{0} \text{ for } t > 0, \quad \mathbf{x}(0) = \mathbf{x}'(0) = \mathbf{0},$$

whose unique solution is the trivial one $\mathbf{x} \equiv \mathbf{0}$. Finally, for Example 4.4, we find the non-linear differential problem

$$-u(x)'' + \frac{\alpha}{2}\frac{1}{u'(x)u''(x)} + g(x) = 0 \text{ in } (0,1), \quad u(0) = 0, u(1) = L > 0.$$

It is not easy to find explicitly the solution $u(x)$ of this problem even for special choices of the function $g(x)$. What we can be sure about, based on strict convexity arguments as we pointed out earlier, is that there is a unique solution.

Example 4.10 We analyze next the classic problem of the brachistochrone for which we are interested in finding the optimal solution, if any, of the variational problem

$$\text{Minimize in } u(x): \quad B(u) = \int_0^A \frac{\sqrt{1+u'(x)^2}}{\sqrt{x}}\,dx$$

subject to

$$u(0) = 0, \quad u(A) = a.$$

We have already insisted in that the direct method cannot be applied in this case due to the linear growth of the integrand

$$F(x, u, U) = \frac{1}{\sqrt{x}}\sqrt{1 + U^2},$$

with respect to the U-variable, though it is strictly convex. We still have the hope to find and show the unique minimizer of the problem through Theorem 4.5 and Proposition 4.3. This requires first to find the solution of the corresponding conditions of optimality, and afterwards, check if hypotheses in Theorem 4.5 hold to conclude.

The Euler-Lagrange equation for this problem reads

$$\frac{d}{dx}\left(\frac{1}{\sqrt{x}}\frac{u'(x)}{\sqrt{1 + u'(x)^2}}\right) = 0 \text{ in } (0, A),$$

or

$$\frac{u'(x)}{\sqrt{x(1 + u'(x)^2)}} = \frac{1}{c} \tag{4.28}$$

for c, a constant. After a bit of algebra, we have

$$u'(x)^2 = \frac{x}{c^2 - x}, \quad u(x) = \int_0^x \sqrt{\frac{s}{c^2 - s}}\, ds.$$

The constant c would be selected by demanding

$$a = \int_0^A \sqrt{\frac{s}{c^2 - s}}\, ds.$$

This form of the solution is not particularly appealing. One can find a better equivalent form by making an attempt to describe the solution in parametric form with the goal of introducing a good change of variable in the integral defining $u(x)$, and calculating it in a more explicit form. Indeed, if we take

$$s(r) = c^2 \sin^2(r/2) = \frac{c^2}{2}(1 - \cos r),$$

and perform the corresponding change of variables in the definition of $u(x)$, we find

$$u(\tau) = c^2 \int_0^\tau \sin^2(\frac{r}{2})\, dr = \frac{c^2}{2}(\tau - \sin\tau),$$

where the upper limit in the integral is

$$x = \frac{c^2}{2}(1 - \cos\tau).$$

We, then, see that the parametric form of the optimal profile is

$$(t(\tau), u(\tau)) = C(1 - \cos\tau, \tau - \sin\tau), \quad \tau \in [0, 1],$$

where the constant C is chosen to ensure that

$$t(1) = A, \quad u(1) = a.$$

This family of curves is well-known in Differential Geometry: they are arcs of cycloids, and enjoy quite remarkable properties.

It is easy to realize that this arc of cycloid $u(x)$ is in fact a Lipschitz function belonging $W^{1,\infty}(0, A)$, and hence it is a function in $W^{1,1}(0, A)$ too. Since the integrand does not depend explicitly upon u, assumptions in Theorem 4.5 and Proposition 4.3 hold true, and we can conclude that this is the only solution of the brachistochrone problem.

One of the most surprising features of this family of curves is its tautochrone condition: the transit time of the mass falling along an arc of a cycloid is independent of the point where the mass starts to fall from !! No matter how high or how low you put the mass, they both will employ the same time interval in getting to the lowest point. This is not hard to argue, once we have gone through the above computations. Notice that the cycloid is parameterized by

$$(t(\tau), u(\tau)) = C(1 - \cos\tau, \tau - \sin\tau)$$

for some constant C and $\tau \in [0, 1]$, so that it is easy to realize that $u'(\tau) = t(\tau)$. If we go back to (4.28), and bear in mind this identity, it is easy to find that indeed

$$\sqrt{\frac{t}{1 + u'(t)^2}} = \tilde{c}, \tag{4.29}$$

a constant. Because in the variable τ the interval of integration for the transit time is constant $[0, 1]$, we see that for the optimal profile, the cycloid, the transit time whose integrand is the inverse of (4.29), is constant independent of the height u_T. This implies the tautochrone feature.

Example 4.11 Optimality conditions for problems in which additional constraints, other than end-point conditions, are to be preserved are more involved. A formal treatment of these is beyond the scope of this text. However, it is not that hard to learn how to deal with them informally in practice. Constraints are dealt with through multipliers, which are additional, auxiliary variables to guarantee that

constraints are taken into account in a fundamental way. There is one multiplier for each restriction, either in the form of equality or inequality. If there are local constraints to be respected at each point x of an interval, then we will have one (or more) multipliers that are functions of x. If, on the contrary, we only have global, integral constraints to be respected then each such constraint will involve just one (or more) unknown number as multiplier.

To show these ideas in a particular example, let us consider the classical problem of the hanging cable, introduced in Chap. 1,

$$\text{Minimize in } u(x): \quad \int_0^H u(x)\sqrt{1+u'(x)^2}\,dx$$

under the constraints

$$u(0)=u(H)=0, \quad L=\int_0^H \sqrt{1+u'(x)^2}\,dx,$$

with $L > H$. Since there is an additional global, integral constraint to be respected, in addition to end-point conditions, we setup the so-called augmented integrand

$$F(u,U)=u\sqrt{1+U^2}+\lambda\sqrt{1+U^2} \tag{4.30}$$

where λ is the unknown multiplier in charge of keeping track of the additional integral restriction. Suppose we look at the variational problem, in spite of not knowing yet the value of λ,

$$\text{Minimize in } u(x): \quad \int_0^H (u(x)+\lambda)\sqrt{1+u'(x)^2}\,dx$$

subject to just $u(0) = u(H) = 0$, through its corresponding Euler-Lagrange equation

$$-\left[(u(x)+\lambda)\frac{u'(x)}{\sqrt{1+u'(x)^2}}\right]'+\sqrt{1+u'(x)^2}=0 \text{ in } (0,H),$$
$$u(0)=u(H)=0.$$

The solution set of this problem will be a one-parameter family $u_\lambda(x)$ of solutions. The one we seek is $u_{\lambda_0}(x)$ so that

$$L=\int_0^H \sqrt{1+u'_{\lambda_0}(x)^2}\,dx.$$

We expect to be able to argue, through Theorem 4.5 and Proposition 4.3, that this function $u_{\lambda_0}(x)$ is the unique minimizer we are searching for.

Because the integrand in (4.30) does not depend on the independent variable, the corresponding Euler-Lagrange equation admits an integration in the form (Exercise 16)

$$(u(x)+\lambda)\sqrt{1+u'(x)^2} - u'(x)(u(x)+\lambda)\frac{u'(x)}{\sqrt{1+u'(x)^2}} = c.$$

Some elementary but careful algebra carries us to

$$u'(x) = \frac{1}{c}\sqrt{(u(x)+\lambda)^2 - c^2},$$

and to

$$\frac{du}{\sqrt{(u(x)+\lambda)^2 - c^2}} = \frac{dx}{c}.$$

A typical hyperbolic trigonometric change of variables leads to

$$u(x) = c\cosh\left(\frac{x}{c}+d\right) - \lambda,$$

for constants c, d and λ. These are determined to match the three numerical constraints of the problem, namely,

$$u(0) = u(H) = 0, \quad L = \int_0^H \sqrt{1+u'(x)^2}\,dx.$$

This adjustment, beyond the precise calculation or approximation, informs us that the shape is that of certain hyperbolic cosine, which in this context is called a catenary. To argue that this catenary is indeed the minimizer of the problem, we note that the interpretation of the problem yields $c > 0$ necessarily, i.e. $u(x) + \lambda_0 > 0$, and the integrand

$$F(u, U) = (u + \lambda_0)\sqrt{1+U^2}$$

is strictly convex in U. However, as a function of the pair (u, U) is not convex, and hence we cannot apply directly Theorem 4.5. The unique solution is, nonetheless, the one we have calculated.

Example 4.12 The case in which no boundary condition whatsoever is imposed in a variational problem is especially interesting. The following corollary is similar to the results presented earlier in the chapter. The only point that deserves some comment is the so-called natural boundary condition (4.33) below that requires a more careful use of the integration-by-parts step. See Exercises 8, 12.

Corollary 4.2

1. *Let*

$$F(x, \mathbf{u}, \mathbf{U}) : J \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \tag{4.31}$$

be continuous in $(\mathbf{u}, \mathbf{U})$ *for a.e.* $x \in J$*, and measurable in* x *for every pair* $(\mathbf{u}, \mathbf{U})$*. Suppose that the following three main conditions hold:*

a. *coercivity and growth: there are positive constants* $C_- \le C_+$ *with*

$$C_-\left(|\mathbf{U}|^2 - 1\right) \le F(x, \mathbf{u}, \mathbf{U}) \le C_+\left(|\mathbf{U}|^2 + 1\right),$$

for $(x, \mathbf{u}, \mathbf{U}) \in J \times \mathbb{R}^n \times \mathbb{R}^n$*;*

b. *convexity: for a.e.* $x \in J$*, and every* $\mathbf{u} \in \mathbb{R}^n$*, we have that*

$$F(x, \mathbf{u}, \cdot) : \mathbb{R}^n \to \mathbb{R}$$

is convex.

Then there are optimal solutions for the variational problem

$$\textit{Minimize in } \mathbf{u} \in H^1(J; \mathbb{R}^n) : \quad I(\mathbf{u}) = \int_J F(x, \mathbf{u}(x), \mathbf{u}'(x))\, dx. \tag{4.32}$$

2. *If, in addition,*

$$|F_{\mathbf{u}}(x, \mathbf{u}, \mathbf{U})| \le C(|\mathbf{U}|^2 + 1), \quad |F_{\mathbf{U}}(x, \mathbf{u}, \mathbf{U})| \le C(|\mathbf{U}| + 1)$$

for some constant C*, and* $\mathbf{u} \in H^1(J; \mathbb{R}^n)$ *is a minimizer of the previous problem, then*

$$F_{\mathbf{U}}(x, \mathbf{u}(x), \mathbf{u}'(x))$$

is absolutely continuous in J*,*

$$F_{\mathbf{u}}(x, \mathbf{u}(x), \mathbf{u}'(x)) - \frac{d}{dx} F_{\mathbf{U}}(x, \mathbf{u}(x), \mathbf{u}'(x)) = \mathbf{0} \textit{ a.e. } x \textit{ in } J,$$

and

$$F_{\mathbf{U}}(x, \mathbf{u}(x), \mathbf{u}'(x))\big|_{x = x_0, x_1} = \mathbf{0}. \tag{4.33}$$

4.8 Non-existence

Readers may have realized that our main existence theorem cannot be applied to examples of variational problems included in the first chapter, not because integrands fail to be convex but because they fail to comply with the coercivity condition for an exponent $p > 1$. Indeed, most of those interesting problems have a linear growth at infinity, so that, at first sight, existence of minimizers is compromised. Indeed, taken for granted that there are no troubles concerning the weak closeness of the set of competing functions or paths, we see that there may be two reasons for the lack of minimizers, depending on whether coercivity or convexity fail.

The simplest example of a variational problem without minimizer because of lack of convexity is due to Bolza. Suppose we wish to

$$\text{Minimize in } u(t): \quad \int_0^1 \left[\frac{1}{2}(|u'(t)| - 1)^2 + \frac{1}{2}u(t)^2\right] dt$$

under $u(0) = u(1) = 0$. The integrand for this problem is the density

$$f(u, U) = \frac{1}{2}(|U| - 1)^2 + \frac{1}{2}u^2,$$

which is not convex in the variable U. Coercivity is not an issue as f has quadratic growth in both variables. The reason for lack of minimizers resides in the fact that the two non-negative contributions to the functional are incompatible with each other: there is no way to reconcile the vanishing of the two terms since either you insist in having $|u'| = 1$; or else, $u = 0$. Both things cannot happen simultaneously. Yet, one can achieve the infimum value $m = 0$ through a minimizing sequence in which slopes ± 1 oscillate faster and faster to lead the term with u^2 to zero. This is essentially the behavior of minimizing sequences due to lack of convexity, and we refer to it as a fine oscillatory process.

A more involved, but similar phenomenon, takes place with the problem in which we try to minimize the functional

$$I(\mathbf{u}) = \int_0^1 \left(|u_1'(x)|\,|u_2'(x)| + |\mathbf{u}(x) - (x, x)|^2\right) dx, \quad \mathbf{u} = (u_1, u_2),$$

in the feasible set of functions

$$\mathcal{A} = \{\mathbf{u} \in H^1((0, 1); \mathbb{R}^2) : \mathbf{u}(0) = \mathbf{0}, \mathbf{u}(1) = (1, 1)\}.$$

It should be noted, however, that lack of convexity does not always lead to lack of minimizers. As a matter of fact, there are some interesting results about existence without convexity. See the final Appendix.

There is also a classical simple example due to Weierstrass for which the lack of uniform coercivity is the reason for lack of minimizer.s Let us look at

$$\text{Minimize in } u(t): \quad \int_{-1}^{1} \frac{1}{2} t^2 u'(t)^2 \, dt$$

under

$$u(-1) = -1, \quad u(1) = 1.$$

We notice that the factor t^2 accounting for coercivity degenerates in the origen. It is precisely this lack of coercivity, even in a single point, that is responsible for the lack of minimizers. If there were some minimizer u, then it would be a solution of the corresponding Euler-Lagrange equation which it is very easily written down as

$$(t^2 u'(t))' = 0 \text{ in } (-1, 1).$$

The general solution of this second-order differential equation is

$$u(t) = \frac{a}{t} + b, \quad a, b, \text{ constants.}$$

The end-point conditions lead to $b = 0$, $a = 1$, and so the minimizer would be

$$u(t) = 1/t$$

which blows up at $t = 0$, and hence it does not belong to any of the spaces $W^{1,p}(0, 1)$.

Example 4.13 The work done by a non-conservative force field. Suppose that we consider the functional

$$W(\mathbf{x}) = \int_0^T \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) \, dt$$

furnishing the work performed by a non-conservative force field $\mathbf{F}(\mathbf{x})$ in going from a point $\mathbf{x}_0$ to $\mathbf{x}_T$ in a time interval $[0, T]$ through a continuous path

$$\mathbf{x}(t) : [0, T] \to \mathbb{R}^n, \quad \mathbf{x}(0) = \mathbf{x}_0, \mathbf{x}(T) = \mathbf{x}_T.$$

It is easy, but surprising, to check that the Euler-Lagrange system is void !! It does not provide any information. This is so regardless of whether the force field is conservative. To better comprehend the situation for a non-trivial, non-conservative force field $\mathbf{F}$, suppose that the system of ODE

$$\mathbf{X}'(t) = -\mathbf{F}(\mathbf{X}(t)), \quad t > 0,$$

has a periodic solution passing through a given point $\mathbf{x}_P \in \mathbb{R}^n$ with period $P > 0$, so that

$$\mathbf{X}(0) = \mathbf{x}_P, \quad \mathbf{X}(P) = \mathbf{x}_P.$$

In this situation, the integral

$$\int_0^P \mathbf{F}(\mathbf{X}(t)) \cdot \mathbf{X}'(t)\, dt = -\int_0^P |\mathbf{F}(\mathbf{X}(t))|^2\, dt$$

is a strictly negative fixed quantity $M < 0$. If a fixed time interval $T > 0$ is given, and take $\lambda = jP/T$ for an arbitrary positive integer j, then the path

$$\mathbf{x}(t) = \mathbf{X}\left(\frac{jP}{T}t\right), \quad t \in [0, T],$$

is such that

$$\mathbf{x}(0) = \mathbf{x}(T) = \mathbf{x}_P,$$

and by an easy change of variables, and periodicity, we find that

$$W(\mathbf{x}) = jM.$$

In this way, we see that if m is the infimum of the work functional under the given conditions, then $m = -\infty$. It suffices to select a sequence of paths $\mathbf{x}_j(t)$ as indicated, reserving an initial time subinterval to go from $\mathbf{x}_0$ to $\mathbf{x}_P$, and another final time subinterval to move back from $\mathbf{x}_P$ to $\mathbf{x}_1$, after going around the periodic solution $\mathbf{X}$ j times. Note that it is imposible for a gradient field to have periodic integral curves as the above curve $\mathbf{X}$.

Is there anything else interesting to do with such non-convex examples? Is there anything else to be learnt from them? We will defer the answer to these appealing questions until Chap. 8 in the context of multidimensional variational problems.

4.9 Exercises

1. Let $\mathbb{H}$ be a vector space, and $I : \mathbb{H} \to \mathbb{R}$, a functional. Show that if I is strictly convex, then it cannot have more than one minimizer, and every local minimizer of I is, in fact, a global minimizer.
2. Let

$$\phi_a(u) = u^4 + au^3 + \frac{3}{8}a^4u^2,$$

where a is a real positive parameter. Study the family of problems

$$\text{Minimize in } u(x): \quad \int_0^1 \phi_a(u'(t))\,dt$$

under end-point conditions

$$u(0) = 0, u(1) = \alpha.$$

We are especially interested in the dependence of minimizers on a and on α.

3. Argue that all minimizers of the functional

$$I(u) = \int_a^b \left[\frac{1}{2}u'(x)^2 + \frac{1}{3}u(x)^3\right] dx$$

are convex functions in (a, b).

4. Provide full details for the necessity part of the proof of Theorem 4.2.
5. An integrand

$$F(\mathbf{x}, \mathbf{z}) : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$$

is a null-lagrangian if the integrals

$$\int_a^b F(\mathbf{x}(t), \mathbf{x}'(t))\,dt$$

only depend upon the starting point $\mathbf{x}_a$ and the final point $\mathbf{x}_b$, and not on the path itself

$$\mathbf{x}(t) : [a, b] \to \mathbb{R}^N$$

as long as

$$\mathbf{x}(a) = \mathbf{x}_a, \quad \mathbf{x}(b) = \mathbf{x}_b.$$

(a) Argue that for a conservative force field

$$\mathbf{F}(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}^N,$$

the integrand

$$F(\mathbf{x}, \mathbf{z}) = \mathbf{F}(\mathbf{x}) \cdot \mathbf{z}$$

is a null-lagrangean. What is the form of the Euler-Lagrange system in this case?

(b) Study the corresponding variational problem for a non-conservative force field.

6. Provide all missing technical details for a full proof of Theorem 4.4.
7. Determine the infimum of the functional

$$\int_0^1 |u'(x)|^\alpha \, dx$$

over the class of Lipschitz functions ($W^{1,\infty}(0, 1)$) with $u(0) = 0$, $u(1) = 1$, in terms of the exponent $\alpha \in \mathbb{R}$.

8. Derive optimality conditions for a variational problem like the ones in Sect. 4.6 under periodic boundary conditions

$$\mathbf{u}(x_0) - \mathbf{u}(x_1) = \mathbf{0}.$$

9. Examine the form of optimality conditions for a variational problem depending explicitly on second derivatives of the form

$$\int_J F(x, \mathbf{u}(x), \mathbf{u}'(x), \mathbf{u}''(x)) \, dx$$

under end-point conditions involving the values of $\mathbf{u}$ and $\mathbf{u}'$ at both end-points of J.

10. The typical way to proceed in the proof of Theorem 4.4 from (4.20) focuses on an integration by parts on the second term rather than on the first. Follow this alternative route, and provide the details to find the same final result.
11. For each positive ϵ, consider the problem that consists in minimizing in pairs $(u_1(t), u_2(t))$ the functional

$$\int_0^1 \left[\frac{1}{2}(u_1(t)u_2(t) - 1)^2 + \frac{\epsilon}{2}(u_1'(t)^2 + u_2'(t)^2) \right] dt$$

under

$$u_1(0) = 2, u_2(0) = 1/2, \quad u_1(1) = -1/2, u_2(1) = -2.$$

(a) Argue that there is an optimal pair $\mathbf{u}_\epsilon = (u_{1,\epsilon}, u_{2,\epsilon})$ for each ϵ.

(b) Examine the corresponding Euler-Lagrange system.

(c) Explore the convergence of such optimal path $\mathbf{u}_\epsilon$ as $\epsilon \searrow 0$.

12. For a regular variational problem for an integral functional of the form

$$I(\mathbf{u}) = \int_0^1 F(\mathbf{u}(t), \mathbf{u}'(t))\,dt,$$

derive optimality conditions for the problem in which only the left-hand side condition $\mathbf{u}(0) = \mathbf{u}_0$ is prescribed but the value $\mathbf{u}(1)$ is free. Pay special attention, precisely, to the optimality condition at this free end-point (natural boundary condition).

13. Explore how to enforce Neumann end-point conditions

$$u'(x_0) = u'_0, \quad u'(x_1) = u'_1$$

for arbitrary numbers u'_0, u'_1 in a variational problem of the particular form

$$\int_J \left(\frac{1}{2}u'(x)^2 + F(x, u(x))\right) dx, \quad J = (x_0, x_1).$$

14. Show that the subset

$$\mathcal{A} = \left\{u(x) \in W^{1,1}(0,1) : u(0) = u(1) = 0, \sqrt{2} = \int_0^1 \sqrt{1 + u'(x)^2}\,dx\right\}$$

is not weakly closed in $W^{1,1}(0, 1)$.

15. Check that the unique solution of problem (4.26) is given by formula (4.27).
16. Show that if the integrand F for a integral functional does not depend on the independent variable x, $F = F(\mathbf{u}, \mathbf{U})$, then the Euler-Lagrange system reduces to

$$F(\mathbf{u}(x), \mathbf{u}'(x)) - \mathbf{u}'(x) \cdot F_{\mathbf{U}}(\mathbf{u}(x), \mathbf{u}'(x)) = \text{constant}.$$

17. For paths

$$\mathbf{x}(t) \in H^1([0, T]; \mathbb{R}^n), \quad \mathbf{x}(0) = \mathbf{x}_0,$$

consider the functional

$$E(\mathbf{x}) = \int_0^T \frac{1}{2}|\mathbf{x}'(t) - \boldsymbol{f}(\mathbf{x}(t))|^2\,dt,$$

where $\boldsymbol{f}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous mapping.

(a) If $\boldsymbol{f}$ is Lipschitz

$$|\boldsymbol{f}(\mathbf{z}_1) - \boldsymbol{f}(\mathbf{z}_2)| \le M|\mathbf{z}_1 - \mathbf{z}_2|, \quad M > 0,$$

show that there is a unique minimizer $\mathbf{X}(t)$ for the problem.

(b) If, in addition, f is C^1, then the unique minimizer $\mathbf{X}$ is the unique solution of the Cauchy problem

$$\mathbf{X}'(t) = f(\mathbf{X}(t)), \quad \mathbf{X}(0) = \mathbf{x}_0.$$

18. Let the C^1 mapping

$$f(\mathbf{x}, \mathbf{y}) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$$

be given complying with the conditions:

(a) uniform Lipschitzianity with respect to $\mathbf{y}$:

$$|f(\mathbf{z}_1, \mathbf{y}) - f(\mathbf{z}_2, \mathbf{y})| \le M|\mathbf{z}_1 - \mathbf{z}_2|$$

with a positive constant M valid for every $\mathbf{y}$;

(b) coercivity with respect to $\mathbf{y}$: for every compact set $\mathbf{K} \subset \mathbb{R}^n$,

$$\lim_{|\mathbf{y}|\to\infty} \sup_{\mathbf{x}\in\mathbf{K}} |f(\mathbf{x}, \mathbf{y})| = \infty.$$

Consider the controllability situation

$$\mathbf{x}'(t) = f(\mathbf{x}(t), \mathbf{y}(t)) \text{ in } [0, T], \quad \mathbf{x}(0) = \mathbf{x}_0, \mathbf{x}(T) = \mathbf{x}_T.$$

It is not clear if, or under what restrictions, there is a solution of this problem.

(a) Study the surjectivity of the mapping

$$\mathbf{y} \in \mathbb{R}^n \mapsto \mathbf{x}(T)$$

where $\mathbf{x}(t)$ is the unique solution of

$$\mathbf{x}'(t) = f(\mathbf{x}(t), \mathbf{y}) \text{ in } [0, T], \mathbf{x}(0) = \mathbf{x}_0,$$

by looking at the variational problem

$$\text{Minimize in } (\mathbf{x}(t), \mathbf{y}) \in H^1([0, T]; \mathbb{R}^n) \times \mathbb{R}^n :$$
$$\int_0^T \frac{1}{2}|\mathbf{x}'(t) - f(\mathbf{x}(t), \mathbf{y})|^2 \, dt$$

subjected to

$$\mathbf{x}(0) = \mathbf{x}_0, \quad \mathbf{x}(T) = \mathbf{x}_T.$$

i. show that there is always an optimal pair

$$(\mathbf{x}, \mathbf{y}) \in H^1([0, T]; \mathbb{R}^n);$$

ii. write optimality conditions.

(b) Do the same for variable paths $\mathbf{y}(t)$.

(c) Examine the particular case of a linear map

$$\boldsymbol{f}(\mathbf{x}, y) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n, \quad \boldsymbol{f}(\mathbf{x}, y) = \mathbf{A}\mathbf{x} + \mathbf{b}y,$$

for constant matrix $\mathbf{A}$ and non-vanishing vector $\mathbf{b}$.

19. For a standard variational problem of the form

$$\int_0^1 F(t, \mathbf{u}(t), \mathbf{u}'(t))\, dt$$

explore general optimality conditions of the form

$$\mathbf{u}(s, t) : (-\epsilon, \epsilon) \times H^1([0, 1]; \mathbb{R}^N) \to \mathbb{R}^N$$

if $\mathbf{u}(0, t)$ is a true minimizer under, say, end-point conditions.

20. For a standard variational problem of the form

$$I(\mathbf{u}) = \int_0^1 F(t, \mathbf{u}(t), \mathbf{u}'(t))\, dt,$$

transform the functional

$$\hat{I}(\psi) = I(\mathbf{u}_\psi)$$

for competing fields of the form

$$\mathbf{u}_\psi(t) = \mathbf{u}(\psi^{-1}(t)),$$

where $\mathbf{u}$ is fixed, while ψ varies in the class of diffeomorphims of the interval $[0, 1]$ leaving invariant both end-points, by performing the change of variables $\psi(s) = t$ in the integral defining $I(\mathbf{u}_\psi)$. Investigate optimality conditions for $\hat{I}(\psi)$.

21. For a variational problem of the form

$$\int_0^1 F(t, \mathbf{u}(t), \mathbf{u}'(t))\, dt$$

under a point-wise constraint

$$\phi(\mathbf{u}(t)) = \mathbf{0} \text{ for all } t \in [0, 1], \quad \phi : \mathbb{R}^N \to \mathbb{R}^n,$$

examine optimality conditions with respect to inner-variations of the form

$$\mathbf{u}(\psi(t)), \quad \psi(t) : [0, 1] \to [0, 1], \psi(0) = 0, \psi(1) = 1.$$

22. For a finite number n, and a finite interval $J = (x_0, x_T) \subset \mathbb{R}$, consider

$$f_i(x, u, v) : J \times \mathbb{R} \times \mathbb{R} \to \mathbb{R},$$
$$\boldsymbol{f}(x, u, v) = (f_i(x, u, v))_i : J \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n,$$

and another function $F : \mathbb{R}^n \to \mathbb{R}$. Explore optimality conditions for the functional

$$I(u) = F\left(\int_J \boldsymbol{f}(x, u(x), u'(x))\,dx\right)$$

subjected to typical end-point conditions

$$u(x_0) = u_0, \quad u(x_T) = u_T,$$

assuming every necessary regularity hypothesis on the functions involved. Look specifically to the case of products and quotients for which $n = 2$, and

$$P(u) = \left(\int_j f_1(x, u(x), u'(x))\,dx\right)\left(\int_j f_2(x, u(x), u'(x))\,dx\right),$$
$$Q(u) = \left(\int_j f_1(x, u(x), u'(x))\,dx\right) \Big/ \left(\int_j f_2(x, u(x), u'(x))\,dx\right).$$

23. If a non-singular symmetric matrix-field

$$\mathbf{A}(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}^{N \times N}$$

determines how to measure distances locally around the point $\mathbf{x} \in \mathbb{R}^N$, then we already described that the distance functional

$$\int_0^1 [\mathbf{x}'(t)^T \mathbf{A}(\mathbf{x}(t))\mathbf{x}'(t)]^{1/2}\,dt$$

provides the length of the curve

$$\mathbf{x}(t) : [0, 1] \to \mathbb{R}^N, \quad \mathbf{x}(0) = \mathbf{P}, \mathbf{x}(1) = \mathbf{Q}$$

joining the two given points. Geodesics are the paths minimizing such functional. Study the corresponding Euler-Lagrange system, and identify the Chrystoffel symbols Γ_{ij}^k of Differential Geometry when the system is written in the form

$$x_k''(t) - \sum_{i,j=1}^{N} \Gamma_{ij}^k (x_i)' x_j' = 0.$$

Start with the case $N = 2$, and try to figure out the general case from there.

24. Consider the functional

$$I(\mathbf{u}) = \int_0^1 |\mathbf{u}'(t)|^2 \, dt, \quad \mathbf{u} : [0, 1] \mapsto \mathbb{R}^N, \mathbf{u} \in H^1(0, 1; \mathbb{R}^N),$$

and define

$$I_i(\mathbf{u}) = \inf_{\psi} I(\mathbf{u}_\psi), \quad \mathbf{u}_\psi = \mathbf{u} \circ \psi,$$

where ψ runs through all diffeomorphisms of the unit interval preserving its two-end points.

(a) For each feasible $\mathbf{u}$ fixed, define

$$\hat{I}(\psi) = I(\mathbf{u}_\psi),$$

and compute the optimal ψ.

(b) Calculate

$$\inf_{\psi} \hat{I}(\psi),$$

for each fixed $\mathbf{u}$. Is $I_i(\mathbf{u})$ a typical integral functional?

(c) Redo the same calculations with the functional

$$I(\mathbf{u}) = \int_0^1 |\mathbf{u}'(t)| \, dt, \quad \mathbf{u} : [0, 1] \mapsto \mathbb{R}^N, \mathbf{u} \in W^{1,1}(0, 1; \mathbb{R}^N).$$

What is the main difference between the two cases?

25. Suppose the family of functions $\{a_\epsilon(x)\}$ is strictly positive and uniformly bounded

$$0 < m \leq a_\epsilon(x) \leq M < +\infty, \quad x \in [0, 1].$$

Consider the variational problem

$$\text{Minimize in } H^1(0,1): \quad Q_\epsilon(u) = \int_0^1 \frac{a_\epsilon(x)}{2} u'(x)^2\,dx$$

subject to $u(x) - x \in H_0^1(0,1)$.

(a) Check that there is a unique minimizer u_ϵ, which is a uniformly bounded sequence in $H^1(0,1)$, and therefore, for a suitable sequence (not relabeled), there some feasible $u \in H^1(0,1)$, $u(x) - x \in H_0^1(0,1)$, such that $u_\epsilon \rightharpoonup u$.
(b) Use optimality to derive a closed form for u_ϵ.
(c) See how to arrange things in such a way that the limit function u becomes the minimizer of a similar variational problem

$$\text{Minimize in } H^1(0,1): \quad Q_\epsilon(u) = \int_0^1 \frac{a(x)}{2} u'(x)^2\,dx$$

under the same end-point conditions.

Part II
Basic Operator Theory

Chapter 5
Continuous Operators

Though we have insisted in the pivotal rule played historically by the Calculus of Variations in the initial steps of Functional Analysis as a discipline on its own, it has grown to be an immense and fruitful field, crucial to many other areas of Mathematics and Science. In particular, Operator Theory is one of those fundamental parts of Functional Analysis that we briefly treat in the following two chapters. We focus on some of the more relevant results that applied analysts and applied mathematicians should know. Most of the concepts and results in this chapter are abstract. But they are so important for fundamental applications in Analysis when spaces become either Lebesgue or Sobolev spaces, as we will see in later chapters, that they need to be taken into account very seriously.

5.1 Preliminaries

We start by considering linear, continuous operators, between Banach spaces which are a natural extension of the ideas discussed earlier in Sect. 2.7 about the dual space.

Proposition 5.1 *Let $\mathbb{E}$, $\mathbb{F}$ be Banach spaces. A linear map $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ is continuous if and only if there is a positive constant M such that*

$$\|\mathbf{T}\mathbf{u}\| \le M\|\mathbf{u}\|, \quad \mathbf{u} \in \mathbb{E}.$$

The proof of this fact is easy and similar to the one in Sect. 2.7. It makes the following definition sensible.

Definition 5.1 If $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ is continuous, we put

$$\|\mathbf{T}\| = \inf\{M > 0 : \|\mathbf{T}\mathbf{u}\| \le M\|\mathbf{u}\|\} = \sup_{\|\mathbf{u}\|=1} \|\mathbf{T}\mathbf{u}\|.$$

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4_5

The space of linear continuous operators from $\mathbb{E}$ to $\mathbb{F}$, written $\mathcal{L}(\mathbb{E}, \mathbb{F})$, becomes a Banach space under this norm. This is straightforward to check.

Note that a bijection $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ is an isomorphism if there is a positive constant $M > 0$ such that

$$\frac{1}{M}\|\mathbf{u}\| \leq \|\mathbf{Tu}\| \leq M\|\mathbf{u}\|, \quad \mathbf{u} \in \mathbb{E}.$$

In particular, two different norms $\|\cdot\|_1$, $\|\cdot\|_2$, in the same underlying space $\mathbb{E}$ are equivalent if there is a constant $M > 0$ with

$$\frac{1}{M}\|\mathbf{u}\|_1 \leq \|\mathbf{u}\|_2 \leq M\|\mathbf{u}\|_1, \quad \mathbf{u} \in \mathbb{E}.$$

Complete metric spaces enjoy quite remarkable properties. Since Banach spaces are a subclass of the class of complete metric spaces, we are interested in stressing some fundamental properties of those spaces related to Baire category methods. There are various equivalent definitions of Baire spaces. We stick to the following.

Definition 5.2 A topological space $\mathbb{X}$ is a Baire space if the countable intersection of open and dense subsets of $\mathbb{X}$ is dense in $\mathbb{X}$. Alternatively, $\mathbb{X}$ is a Baire space if for every sequence of closed subsets with empty interior, the union of them all still has empty interior.

The fundamental result we are interested in is the fact that every complete metric space is a Baire space.

Theorem 5.1 *Every complete metric space is a Baire space.*

Proof Let $\mathbb{X}$ be a complete metric space with distance function $d(\cdot, \cdot)$, and let $\mathbf{O}_j$ be open and dense. We pretend to show that $\mathbf{O} = \cap_j \mathbf{O}_j$ is dense too. To this end, choose an additional arbitrary open, non-empty set $\mathbf{G}$, and show that $\mathbf{G} \cap \mathbf{O}$ is non-empty.

Take some $\mathbf{x}_0 \in \mathbf{G}$, and, given that $\mathbf{G}$ is open, $r_0 > 0$ so that

$$\mathbf{B}_{r_0}(\mathbf{x}_0) \equiv \{\mathbf{x} \in \mathbb{X} : d(\mathbf{x}, \mathbf{x}_0) \leq r_0\} \subset \mathbf{G}.$$

Since $\mathbf{O}_1$ is open and dense, we can find $\mathbf{x}_1$ and $r_1 > 0$ with

$$\mathbf{B}_{r_1}(\mathbf{x}_1) \equiv \{\mathbf{x} \in \mathbb{X} : d(\mathbf{x}, \mathbf{x}_1) \leq r_1\} \subset \mathbf{B}_{r_0}(\mathbf{x}_0) \cap \mathbf{O}_1, \quad 0 < r_1 < r_0/2.$$

In a recursive way, we find sequences $\{\mathbf{x}_j\}$ and $\{r_j\}$ such that

$$\mathbf{B}_{r_{j+1}}(\mathbf{x}_{j+1}) \subset \mathbf{B}_{r_j}(\mathbf{x}_j) \cap \mathbf{O}_{j+1}, \quad r_{j+1} < r_j/2.$$

It is then clear that $\{\mathbf{x}_j\}$ is a Cauchy sequence because for all j,

$$d(\mathbf{x}_{j+1} - \mathbf{x}_j) \leq r_j,$$

and, hence, if $j < k$,

$$d(\mathbf{x}_k, \mathbf{x}_j) \leq \sum_{i=j}^{k-1} d(\mathbf{x}_i, \mathbf{x}_{i+1}) \leq \sum_{i=j}^{k-1} r_i \leq r_j \sum_{i=0}^{k-1-j} 2^{-i} = r_j 2^{j-k+1}.$$

The completeness of the space leads to a limit $\mathbf{x}$. Since the sequence of balls $\{\mathbf{B}_{r_j}(\mathbf{x}_j)\}$ is a nested, decreasing sequence of subsets of $\mathbf{G} \cap \mathbf{O}_{j-1}$, the limit $\mathbf{x}$ will belong to $\mathbf{G} \cap \mathbf{O}$. This proves the result. □

5.2 The Banach-Steinhaus Principle

The Banach-Steinhaus principle enables to derive a uniform bound from a pointwise bound, which is a quite remarkable fact given that, in general, it is far from being so.

Theorem 5.2 *Let* $\mathbb{E}$ *and* $\mathbb{F}$ *be two Banach spaces, and* $\{\mathbf{T}_j\}_{j \in J}$*, a family (not necessarily a sequence) in* $\mathcal{L}(\mathbb{E}, \mathbb{F})$*. The following two assertions are equivalent:*

1. there is a constant $C > 0$ *such that*

$$\|\mathbf{T}_j \mathbf{x}\| \leq C \|\mathbf{x}\|, \quad \mathbf{x} \in \mathbb{E}, j \in J;$$

2.

$$\sup_{j \in J} \|\mathbf{T}_j \mathbf{x}\| < \infty, \quad \mathbf{x} \in \mathbb{E}. \tag{5.1}$$

Proof Suppose (5.1) holds, and set, for each $i \in \mathbb{N}$,

$$\mathbf{F}_i = \{\mathbf{x} \in \mathbb{E} : \|\mathbf{T}_j \mathbf{x}\| \leq i, \text{ for all } j \in J\}.$$

It is clear that each $\mathbf{F}_i$ is closed, and $\mathbb{E} = \cup_i \mathbf{F}_i$, precisely by (5.1). Since every Banach space is a complete metric space, by Theorem 5.1 and Definition 5.2, there must be some i_0 with $\mathbf{F}_{i_0}$ having a non-empty interior. This means that there is some $\mathbf{x}_0 \in \mathbb{E}$ and $\rho > 0$ with

$$\|\mathbf{T}_j(\mathbf{x}_0 + \rho \mathbf{x})\| \leq i_0, \quad j \in J, \|\mathbf{x}\| \leq 1.$$

This inequality can be recast as

$$\|\mathbf{T}_j\| \leq \frac{1}{\rho}(i_0 + \|\mathbf{T}_j \mathbf{x}_0\|)$$

which is our conclusion. □

An interesting corollary is the following.

Corollary 5.1 *Let* $\mathbb{E}$, $\mathbb{F}$ *be two Banach spaces, and* $\{\mathbf{T}_j\}$ *a sequence in* $\mathcal{L}(\mathbb{E}, \mathbb{F})$, *such that the limit of* $\{\mathbf{T}_j \mathbf{x}\}$ *exists for every* $\mathbf{x} \in \mathbb{E}$, *and is denoted by* $\mathbf{Tx}$. *Then* $\mathbf{T} \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ *and*

$$\|\mathbf{T}\| \leq \liminf_{j \to \infty} \|\mathbf{T}_j\|.$$

Proof Directly from Theorem 5.2, we conclude that $\mathbf{T} \in \mathcal{L}(\mathbb{E}, \mathbb{F})$. Since, it is true that

$$\|\mathbf{T}_j \mathbf{x}\| \leq \|\mathbf{T}_j\| \, \|\mathbf{x}\|$$

for every j, we can conclude the inequality in the statement. □

Remark 5.1 It is important, however, to realize that it is not true, in general, that

$$\|\mathbf{T}_j - \mathbf{T}\| \to 0.$$

5.3 The Open Mapping and Closed Graph Theorems

These two are among the classical results in Functional Analysis that ought to be studied in a first course.

Theorem 5.3 *Every surjective, linear operator* $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ *between two Banach spaces* $\mathbb{E}$ *and* $\mathbb{F}$ *is open, i.e. the image* $\mathbf{T}(\mathbf{O})$ *of every open set* $\mathbf{O}$ *in* $\mathbb{E}$ *is open in* $\mathbb{F}$.

Proof It is easy to argue that it suffices to show that there is some $\rho > 0$ such that

$$\mathbf{B}_\rho \subset \mathbf{T}(\mathbf{B}_1), \tag{5.2}$$

where the first ball is a ball in $\mathbb{F}$ centered at $\mathbf{0}$, while the second ball is the unit ball in $\mathbb{E}$ (exercise). To show (5.2), put

$$\mathbf{F}_i = i\overline{\mathbf{T}(\mathbf{B}_1)}, \quad i = 1, 2, \dots$$

Since $\mathbf{T}$ is onto, it is true that $\mathbb{F} = \cup_i \mathbf{F}_i$. By the Baire category result Theorem 5.1, there must be some i_0 with $\mathbf{F}_{i_0}$ not having non-empty interior, for otherwise the union would have empty interior which is not the case. There are, therefore, $\rho > 0$ and $\mathbf{y}_0 \in \mathbb{F}$ with

$$\mathbf{B}_{4\rho}(\mathbf{y}_0) \subset \overline{\mathbf{T}(\mathbf{B}_1)}.$$

Due to the symmetry of $\overline{\mathbf{T}(\mathbf{B}_1)}$ with respect to changes of sign, it is easy to have

$$\mathbf{B}_{4\rho}(\mathbf{0}) \subset \overline{\mathbf{T}(\mathbf{B}_1)} + \overline{\mathbf{T}(\mathbf{B}_1)} = 2\overline{\mathbf{T}(\mathbf{B}_1)},$$

or

$$\mathbf{B}_{2\rho}(\mathbf{0}) \subset \overline{\mathbf{T}(\mathbf{B}_1)}. \tag{5.3}$$

Based on (5.3), we would like to prove that

$$\mathbf{B}_\rho \equiv \mathbf{B}_\rho(\mathbf{0}) \subset \mathbf{T}(\mathbf{B}_1).$$

To this end, let $\mathbf{y} \in \mathbb{F}$ be such that $\|\mathbf{y}\| \leq \rho$; we want to find $\mathbf{x} \in \mathbb{E}$ with

$$\|\mathbf{x}\| < 1, \quad \mathbf{Tx} = \mathbf{y}.$$

Because of (5.3), for $\epsilon > 0$ arbitrary, a vector $\mathbf{z} \in \mathbb{E}$ can be found so that

$$\|\mathbf{z}\| < 1/2, \quad \|\mathbf{y} - \mathbf{Tz}\| < \epsilon.$$

In particular, for $\epsilon = \rho/2$, there is some $\mathbf{z}_1 \in \mathbb{E}$ such that

$$\|\mathbf{z}_1\| < 1/2, \quad \|\mathbf{y} - \mathbf{Tz}_1\| < \rho/2.$$

But this same argument applied to the difference $\mathbf{y} - \mathbf{Tz}_1$, instead of $\mathbf{y}$, and for $\epsilon = \rho/4$, yields $\mathbf{z}_2 \in \mathbb{E}$ and

$$\|\mathbf{z}_2\| < 1/4, \quad \|(\mathbf{y} - \mathbf{Tz}_1) - \mathbf{Tz}_2\| < \rho/4.$$

Proceeding recursively, we would find a sequence $\{\mathbf{z}_i\}$ with the properties

$$\|\mathbf{z}_i\| < 2^{-i}, \quad \left\| \mathbf{y} - \mathbf{T}\left(\sum_{k=1}^{i} \mathbf{z}_k\right) \right\| < \rho 2^{-i},$$

for every i. Consequently, the sequence of partial sums

$$\mathbf{x}_i = \sum_{k=1}^{i} \mathbf{z}_k$$

is a Cauchy sequence, hence convergent to some $\mathbf{x} \in \mathbf{B}_1$, and, by continuity of $\mathbf{T}$,

$$\mathbf{T}\mathbf{x} = \mathbf{y},$$

as desired. □

There are at least three important consequences of this fundamental result, one of which is the closed graph theorem which we discuss in the third place.

Corollary 5.2 *If* $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ *is a linear, continuous operator between two Banach spaces that is a bijection, then* $\mathbf{T}^{-1} : \mathbb{F} \to \mathbb{E}$ *is linear, and continuous too.*

Proof We investigate the consequences of injectivity together with property (5.2). Indeed, if $\mathbf{T}$ is one-to-one, and $\mathbf{x}$ is such that $\|\mathbf{T}\mathbf{x}\| < \rho$ then necessarily $\|\mathbf{x}\| < 1$. For an arbitrary $\mathbf{x} \in \mathbb{E}$, the vector

$$\overline{\mathbf{x}} = \frac{\rho}{2\|\mathbf{T}\mathbf{x}\|}\mathbf{x}$$

is such that $\|\mathbf{T}\overline{\mathbf{x}}\| < \rho$, and hence

$$\|\mathbf{x}\| < \frac{\rho}{2}\|\mathbf{T}\mathbf{x}\|.$$

This is the continuity of $\mathbf{T}^{-1}$. □

A particularly interesting situation of this kind occurs when a unique underlying space turns out to be a Banach space under two different norms.

Corollary 5.3 *If a vector space* $\mathbb{E}$ *is a Banach space under two different norms* $\|\cdot\|_1$ *and* $\|\cdot\|_2$*, and there is a constant* $C > 0$ *with*

$$\|\mathbf{x}\|_2 \le C\|\mathbf{x}\|_1, \quad \mathbf{x} \in \mathbb{E},$$

then the two norms are equivalent, i.e. there is a further constant $c > 0$ *with*

$$\|\mathbf{x}\|_1 \le c\|\mathbf{x}\|_2, \quad \mathbf{x} \in \mathbb{E}.$$

Proof The proof is hardly in need of any clarification. Simply apply the previous corollary to the identity as a continuous bijection between the two Banach spaces $(\mathbb{E}, \|\cdot\|_i)$, $i = 1, 2$. □

It is of paramount importance that $\mathbb{E}$ is a Banach space under the two norms.

Finally, we have the closed graph theorem. For an operator $\mathbf{T} : \mathbb{E} \to \mathbb{F}$, its graph $\mathbf{G}(\mathbf{T})$ is the subset of the product space given by

$$\mathbf{G}(\mathbf{T}) = \{(\mathbf{x}, \mathbf{T}\mathbf{x}) \subset \mathbb{E} \times \mathbb{F} : \mathbf{x} \in \mathbb{E}\}.$$

Theorem 5.4 *Let* $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ *be a linear operator between two Banach spaces. The two following assertions are equivalent:*

1. $\mathbf{T}$ *is continuous;*
2. $\mathbf{G}(\mathbf{T})$ *is closed in* $\mathbb{E} \times \mathbb{F}$.

Proof If $\mathbf{T}$ is continuous then its graph is closed. This is even true for non-linear operators. For the converse, consider in $\mathbb{E}$ the graph norm

$$\|\mathbf{x}\|_0 = \|\mathbf{x}\|_{\mathbb{E}} + \|\mathbf{T}\mathbf{x}\|_{\mathbb{F}}.$$

Relying on the assumption of the closedness of the graph, it is easy to check that $\mathbb{E}$ is also a Banach space under $\|\cdot\|_0$ (exercise). We trivially have

$$\|\mathbf{x}\|_{\mathbb{E}} \leq \|\mathbf{x}\|_0, \quad \mathbf{x} \in \mathbb{E}.$$

By the preceding corollary, the two norms are equivalent

$$\|\mathbf{x}\|_0 \leq c\|\mathbf{x}\|_{\mathbb{E}}, \quad \mathbf{x} \in \mathbb{E},$$

for some positive constant c. But then

$$\|\mathbf{T}\mathbf{x}\|_{\mathbb{F}} \leq c\|\mathbf{x}\|_{\mathbb{E}}, \quad \mathbf{x} \in \mathbb{E},$$

and $\mathbf{T}$ is continuous. □

5.4 Adjoint Operators

There is a canonical operation associated with a linear, continuous mapping $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ between two Banach spaces with respective dual spaces $\mathbb{E}'$, $\mathbb{F}'$. If $T \in \mathbb{F}'$ so that $T : \mathbb{F} \to \mathbb{R}$ is linear and continuous, the composition

$$T \circ \mathbf{T} : \mathbb{E} \to \mathbb{R}$$

is linear and continuous as well, and hence it belongs to $\mathbb{E}'$.

Definition 5.3 Let $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ be linear and continuous. The map $\mathbf{T}' : \mathbb{F}' \to \mathbb{E}'$ defined through the identity

$$\langle \mathbf{T}'T, \mathbf{x} \rangle = \langle T, \mathbf{T}\mathbf{x} \rangle, \quad \mathbf{x} \in \mathbb{E},$$

is called the adjoint operator of $\mathbf{T}$.

It is very easy to check that

$$\|\mathbf{T}'\|_{\mathcal{L}(\mathbb{F}',\mathbb{E}')} = \|\mathbf{T}\|_{\mathcal{L}(\mathbb{E},\mathbb{F})}.$$

The duality between a Banach space $\mathbb{E}$ and its dual $\mathbb{E}'$ can be understood as a bilinear mapping

$$\langle \cdot, \cdot \rangle : \mathbb{E} \times \mathbb{E}' \to \mathbb{R}.$$

Definition 5.4

1. For a subspace $\mathbb{M}$ of a Banach space $\mathbb{E}$, we define its orthogonal subspace

$$\mathbb{M}^{\perp} = \{T \in \mathbb{E}' : \langle T, \mathbf{x} \rangle = 0 \text{ for all } \mathbf{x} \in \mathbb{M}\}.$$

   Similarly, if $\mathbb{M}$ is now a subspace of the dual $\mathbb{E}'$, then

$$\mathbb{M}^{\perp} = \{\mathbf{x} \in \mathbb{E} : \langle T, \mathbf{x} \rangle = 0 \text{ for all } T \in \mathbb{M}\}.$$

2. Given a linear, continuous operator $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ there are two special subspaces, one in $\mathbb{E}$, and another one in $\mathbb{F}$, associated with $\mathbf{T}$

$$\mathbb{N}(\mathbf{T}) = \{\mathbf{x} \in \mathbb{E} : \mathbf{T}\mathbf{x} = \mathbf{0}\}, \quad \mathbb{R}(\mathbf{T}) = \{\mathbf{T}\mathbf{x} : \mathbf{x} \in \mathbb{E}\}.$$

It is instructive to relate the various concepts in this definition. To do this in an efficient way, we first show the following which is also a nice fact on its own right.

Proposition 5.2 *If $\mathbb{M}$ is a subspace of a Banach space $\mathbb{E}$, then*

$$\left(\mathbb{M}^{\perp}\right)^{\perp} = \overline{\mathbb{M}}. \tag{5.4}$$

However, if $\mathbb{M}$ is a subspace of the dual $\mathbb{E}'$, then

$$\left(\mathbb{M}^{\perp}\right)^{\perp} \supset \overline{\mathbb{M}}. \tag{5.5}$$

If $\mathbb{E}$ is reflexive, (5.5) becomes an equality too.

Proof If $\mathbb{M}$ is a subspace of the Banach space $\mathbb{E}$, it is easy to check that $\mathbb{M} \subset (\mathbb{M}^\perp)^\perp$, and that $(\mathbb{M}^\perp)^\perp$ is closed. In particular, $(\mathbb{M}^\perp)^\perp$ is a Banach space itself (under the same norm in $\mathbb{E}$). According to Corollary 3.3, to show (5.4) it suffices to prove that for a linear functional $T \in \mathbb{E}'$ such that

$$\langle T, \mathbf{x}\rangle = 0 \text{ for all } \mathbf{x} \in \mathbb{M} \tag{5.6}$$

we have

$$T|_{(\mathbb{M}^\perp)^\perp} \equiv 0. \tag{5.7}$$

But (5.6) exactly means that $T \in \mathbb{M}^\perp$, and so (5.7) holds by definition.

The proof of (5.5) is similar. However, equality might not hold because linear continuous functionals over $\mathbb{E}'$, i.e. elements of $\mathbb{E}''$, vanishing on $\mathbb{M}$ may not belong to $\mathbb{E}$ and not be elements of $\mathbb{M}^\perp$. If $\mathbb{E}$ is assumed to be reflexive, the equality is recovered as before. □

Proposition 5.3 *Let* $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ *be a linear, continuous operator between two Banach spaces, and* $\mathbf{T}' : \mathbb{F}' \to \mathbb{E}'$*, its adjoint. Then*

$$\mathbb{N}(\mathbf{T}) = \mathbb{R}(\mathbf{T}')^\perp, \quad \mathbb{N}(\mathbf{T}') = \mathbb{R}(\mathbf{T})^\perp,$$
$$\mathbb{N}(\mathbf{T})^\perp \supset \overline{\mathbb{R}(\mathbf{T}')}, \quad \mathbb{N}(\mathbf{T}')^\perp = \overline{\mathbb{R}(\mathbf{T})}.$$

If $\mathbb{E}$ *is reflexive, then*

$$\mathbb{N}(\mathbf{T})^\perp = \overline{\mathbb{R}(\mathbf{T}')}.$$

Proof The first two identities are essentially a consequence of the definitions. Concerning the other two, they are implied by the first two if we bear in mind Proposition 5.2, including the situation when $\mathbb{E}$ is reflexive. □

We wrap this section by commenting briefly on the double adjoint $\mathbf{T}'' : \mathbb{E}'' \to \mathbb{F}''$. If we start with $\mathbf{T} : \mathbb{E} \to \mathbb{F}$, we know that the adjoint operator $\mathbf{T}' : \mathbb{F}' \to \mathbb{E}'$ is determined through

$$\langle \mathbf{T}'T, \mathbf{x}\rangle = \langle T, \mathbf{T}\mathbf{x}\rangle, \quad \mathbf{x} \in \mathbb{E}, T \in \mathbb{F}'.$$

Once we have $\mathbf{T}' : \mathbb{F}' \to \mathbb{E}'$, we can talk about $\mathbf{T}'' : \mathbb{E}'' \to \mathbb{F}''$ characterized by

$$\langle \mathbf{T}''\mathbf{x}, T\rangle = \langle \mathbf{x}, \mathbf{T}'T\rangle, \quad \mathbf{x} \in \mathbb{E}'', T \in \mathbb{F}'.$$

In particular, since $\mathbb{E} \subset \mathbb{E}''$ is a closed subspace, if we take $\mathbf{x} \in \mathbb{E}$ we would have for every $T \in \mathbb{F}'$,

$$\langle \mathbf{T}''\mathbf{x}, T\rangle = \langle T, \mathbf{T}\mathbf{x}\rangle.$$

This identity yields, again because $\mathbb{F}$ is a closed subspace of $\mathbb{F}$, that $\mathbf{T}''\mathbf{x} = \mathbf{T}\mathbf{x}$. This fact can be interpreted by saying that $\mathbf{T}'' : \mathbb{E}'' \to \mathbb{F}''$ is an extension of $\mathbf{T} : \mathbb{E} \to \mathbb{F}$. If $\mathbb{E}$ is reflexive, $\mathbf{T}''$ and $\mathbf{T}$ is the same operator with a different target space if $\mathbb{F}$ is not reflexive.

5.5 Spectral Concepts

For a Banach space $\mathbb{E}$, we denote $\mathcal{L}(\mathbb{E})$ the collection of linear, continuous operators from $\mathbb{E}$ to itself. $\mathbf{1} \in \mathcal{L}(\mathbb{E})$ stands for the identity map.

Definition 5.5 Let $\mathbf{T} \in \mathcal{L}(\mathbb{E})$ for a Banach space $\mathbb{E}$.

1. The resolvent $\rho(\mathbf{T})$ of $\mathbf{T}$ is the subset of real numbers λ such that $\mathbf{T} - \lambda\mathbf{1}$ is a bijection from $\mathbb{E}$ onto $\mathbb{E}$. We put

$$\mathbf{R}_\lambda = (\mathbf{T} - \lambda\mathbf{1})^{-1}, \quad \lambda \in \rho(\mathbf{T}).$$

2. The spectrum $\sigma(\mathbf{T})$ is defined to be

$$\sigma(\mathbf{T}) = \mathbb{R} \setminus \rho(\mathbf{T}).$$

   The spectral radius $\nu(\mathbf{T})$ is the number

$$\nu(\mathbf{T}) = \sup\{|\lambda| : \lambda \in \sigma(\mathbf{T})\}.$$

3. A real number $\lambda \in \mathbb{R}$ is an eigenvalue of $\mathbf{T}$ if

$$\mathbb{N}(\mathbf{T} - \lambda\mathbf{1}) \neq \{\mathbf{0}\}.$$

   If λ is an eigenvalue of $\mathbf{T}$, the non-trivial subspace $\mathbb{N}(\mathbf{T} - \lambda\mathbf{1})$ is, precisely, its eigenspace. The set of all eigenvalues of $\mathbf{T}$ is denoted $e(\mathbf{T})$.

Note that we always have

$$e(\mathbf{T}) \subset \sigma(\mathbf{T}).$$

Some basic properties follow. Let $\mathbf{T} \in \mathcal{L}(\mathbb{E})$ for a certain Banach space $\mathbb{E}$.

Proposition 5.4

1. *The resolvent $\rho(\mathbf{T})$ is open in $\mathbb{R}$.*
2. *For every pair $\lambda, \mu \in \rho(\mathbf{T})$,*

$$\mathbf{R}_\lambda - \mathbf{R}_\mu = (\lambda - \mu)\mathbf{R}_\mu\mathbf{R}_\lambda. \tag{5.8}$$

In particular,

$$\lim_{\lambda \to \mu} \frac{\mathbf{R}_\lambda - \mathbf{R}_\mu}{\lambda - \mu} = \mathbf{R}_\mu^2.$$

3. *The spectrum* $\sigma(\mathbf{T})$ *is a compact set of* $\mathbb{R}$*.*
4. *The spectral radius* $\nu(\mathbf{T})$ *is smaller than the norm* $\|\mathbf{T}\|$*.*

Proof We first check that the class of isomorphisms in $\mathcal{L}(\mathbb{E})$ is open in this Banach space. Indeed, if $\mathbf{T} \in \mathcal{L}(\mathbb{E})$ is an isomorphism, the ball

$$\mathbb{B} = \{\mathbf{S} \in \mathcal{L}(\mathbb{E}) : \|\mathbf{S} - \mathbf{T}\| < 1/\|\mathbf{T}^{-1}\|\}$$

is fully contained in the same class. To check this, we can write

$$\mathbf{S} = \mathbf{T} - (\mathbf{T} - \mathbf{S}) = \mathbf{T}\left(\mathbf{1} - \mathbf{T}^{-1}(\mathbf{T} - \mathbf{S})\right). \tag{5.9}$$

But if $\mathbf{S} \in \mathbb{B}$,

$$\|\mathbf{T}^{-1}(\mathbf{T} - \mathbf{S})\| < 1,$$

and then (see exercise below)

$$\mathbf{1} - \mathbf{T}^{-1}(\mathbf{T} - \mathbf{S})$$

is an isomorphism. Equation (5.9) shows that so is $\mathbf{S}$. Since the resolvent of $\mathbf{T}$ is the pre-image of the class of isomorphisms, an open subset in $\mathcal{L}(\mathbb{E})$ by our argument above, through the continuous map

$$\psi(\lambda) : \mathbb{R} \to \mathcal{L}(\mathbb{E}), \quad \psi(\lambda) = \mathbf{T} - \lambda\mathbf{1},$$

we can conclude that the resolvent is an open set of $\mathbb{R}$.

Formula (5.8) is straightforward. In fact

$$\begin{aligned} \mathbf{R}_\lambda - \mathbf{R}_\mu =& \mathbf{R}_\mu(\mathbf{R}_\mu^{-1}\mathbf{R}_\lambda - \mathbf{1}) \\ =& \mathbf{R}_\mu(\mathbf{R}_\mu^{-1} - \mathbf{R}_\lambda^{-1})\mathbf{R}_\lambda \\ =& (\lambda - \mu)\mathbf{R}_\mu\mathbf{R}_\lambda. \end{aligned}$$

The spectrum is a closed set of $\mathbb{R}$, given that its complement, the resolvent, has been shown to be open. Moreover, it is easy to prove that

$$\sigma(\mathbf{T}) \subset [-\|\mathbf{T}\|, \|\mathbf{T}\|], \tag{5.10}$$

which, by the way, proves the last statement of the proposition. Equation (5.10) is a direct consequence of the exercise mentioned above: $\|(1/\lambda)\mathbf{T}\| < 1$ implies that $\mathbf{1} - (1/\lambda)\mathbf{T}$ is an isomorphism, and so is $\mathbf{T} - \lambda\mathbf{1}$, i.e. $\lambda \in \rho(\mathbf{T})$. If $|\lambda| > \|\mathbf{T}\|$ what we have just argued is correct. □

5.6 Self-Adjoint Operators

Since the dual space $\mathbb{H}'$ of a Hilbert space $\mathbb{H}$ can be identified canonically with itself $\mathbb{H}' = \mathbb{H}$, given an operator $\mathbf{T} \in \mathcal{L}(\mathbb{H})$, its adjoint operator $\mathbf{T}'$ can also be regarded as a member of $\mathcal{L}(\mathbb{H})$, and is determined through the condition

$$\langle \mathbf{Tx}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{T}'\mathbf{y}\rangle, \quad \mathbf{x}, \mathbf{y} \in \mathbb{H}.$$

Definition 5.6 Let $\mathbf{T} \in \mathcal{L}(\mathbb{H})$ for a Hilbert space $\mathbb{H}$. $\mathbf{T}$ is self-adjoint if $\mathbf{T}' = \mathbf{T}$, or, equivalently, if

$$\langle \mathbf{Tx}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{Ty}\rangle, \quad \mathbf{x}, \mathbf{y} \in \mathbb{H}.$$

In addition, a self-adjoint operator $\mathbf{T}$ is positive if

$$\langle \mathbf{Tx}, \mathbf{x}\rangle \geq 0, \quad \mathbf{x} \in \mathbb{H}.$$

Theorem 5.5 *For a self-adjoint operator* $\mathbf{T} \in \mathcal{L}(\mathbb{H})$*,*

$$\sigma(\mathbf{T}) \subset [m, M], \quad m = \inf_{\|\mathbf{x}\|=1} \langle \mathbf{Tx}, \mathbf{x}\rangle, M = \sup_{\|\mathbf{x}\|=1} \langle \mathbf{Tx}, \mathbf{x}\rangle.$$

Moreover, both end-points m and M belong to $\sigma(\mathbf{T})$.

It is very easy to argue that $e(\mathbf{T})$ is a subset of the interval $[m, M]$. Indeed, if there is a non-null vector $\mathbf{x}$ such that

$$(\mathbf{T} - \lambda\mathbf{1})\mathbf{x} = \mathbf{0}, \quad \mathbf{Tx} = \lambda\mathbf{x},$$

which can be assumed to have size one $\|\mathbf{x}\| = 1$, we would trivially have

$$\langle \mathbf{Tx}, \mathbf{x}\rangle = \lambda\|\mathbf{x}\|^2 = \lambda,$$

and hence $\lambda \in [m, M]$. However, given that $e(\mathbf{T}) \subset \sigma(\mathbf{T})$, what this proposition ensures is more than that. But it requires the self-adjointness of the operator $\mathbf{T}$. For self-adjoint operators $\mathbf{T}$, it is true that

$$e(\mathbf{T}) \subset \sigma(\mathbf{T}) \subset [m, M].$$

Proof The main tool for the proof is the Lax-Milgram lemma Theorem 3.1. Assume that $\lambda < m$. By our above remark that λ cannot be an eigenvalue, $\mathbf{T} - \lambda\mathbf{1}$ is injective. Take $\epsilon > 0$, so that $m - \lambda \geq \epsilon$. In this case the bilinear form

$$A(\mathbf{x}, \mathbf{y}) = \langle \mathbf{T}\mathbf{x} - \lambda\mathbf{x}, \mathbf{y}\rangle$$

turns out to be symmetric (because $\mathbf{T}$ is self-adjoint), continuous, and coercive, since the associated quadratic form

$$\langle \mathbf{T}\mathbf{x} - \lambda\mathbf{x}, \mathbf{x}\rangle \geq \epsilon\|\mathbf{x}\|^2$$

is coercive. By the Lax-Milgram lemma, for every given element $\mathbf{y} \in \mathbb{H}$, there is $\overline{\mathbf{x}}$ with

$$\langle \mathbf{T}\overline{\mathbf{x}} - \lambda\overline{\mathbf{x}}, \mathbf{z}\rangle = \langle \mathbf{y}, \mathbf{z}\rangle$$

for every $\mathbf{z} \in \mathbb{H}$, i.e.

$$\mathbf{T}\overline{\mathbf{x}} - \lambda\overline{\mathbf{x}} = \mathbf{y},$$

and $\mathbf{T} - \lambda\mathbf{1}$ is onto. This means that $\lambda \in \rho(\mathbf{T})$.

If $\lambda > M$, argue in the same way with the bilinear form

$$A(\mathbf{x}, \mathbf{y}) = \langle \lambda\mathbf{x} - \mathbf{T}\mathbf{x}, \mathbf{y}\rangle.$$

For the last part of the proof, suppose, seeking a contradiction, that $m \in \rho(\mathbf{T})$, so that $\mathbf{T} - m\mathbf{1}$ is a bijection and $(\mathbf{T} - m\mathbf{1})^{-1}$ is continuous. By definition of m, there is a sequence of unit vectors $\mathbf{x}_j$, $\|\mathbf{x}_j\| = 1$, such that

$$\langle \mathbf{T}\mathbf{x}_j - m\mathbf{x}_j, \mathbf{x}_j\rangle \to 0. \tag{5.11}$$

If we could derive from this convergence that the only way to have it is

$$\mathbf{T}\mathbf{x}_j - m\mathbf{x}_j \to 0, \tag{5.12}$$

that would be the contradiction with the continuity of $(\mathbf{T} - m\mathbf{1})^{-1}$ and the fact that the $\mathbf{x}_j$'s are unit vectors.

Concluding (5.12) from (5.11) involves the non-negativity of the bilinear, possibly degenerate, form

$$A(\mathbf{x}, \mathbf{y}) = \langle \mathbf{T}\mathbf{x} - m\mathbf{x}, \mathbf{y}\rangle.$$

This is nothing but the classical Cauchy-Schwarz inequality that does not require the strict positivity of the underlying bilinear form to be correct. We include it next for the sake of completeness and clarity. □

Lemma 5.1 *Let $A(\mathbf{u}, \mathbf{v})$ a symmetric, bilinear, non-negative form in a Hilbert space $\mathbb{H}$. Then*

$$A(\mathbf{u}, \mathbf{v})^2 \geq A(\mathbf{u}, \mathbf{u})A(\mathbf{v}, \mathbf{v}), \quad \mathbf{u}, \mathbf{v} \in \mathbb{H}.$$

Proof The proof is elementary. For an arbitrary pair $\mathbf{u}, \mathbf{v} \in \mathbb{H}$, consider the quadratic, non-negative polynomial

$$P(t) = A(\mathbf{u} + t\mathbf{v}, \mathbf{u} + t\mathbf{v}),$$

and conclude by demanding the non-negativity of the discriminant. □

If we use this lemma for the usual bilinear form

$$A(\mathbf{x}, \mathbf{y}) = \langle \mathbf{T}\mathbf{x} - m\mathbf{x}, \mathbf{y} \rangle,$$

in the context of the convergences (5.12) and (5.11), we see that

$$\|\mathbf{T}\mathbf{x}_j - m\mathbf{x}_j\|^2 \leq \langle \mathbf{T}\mathbf{x}_j - m\mathbf{x}_j, \mathbf{x}_j \rangle \to 0,$$

and the proof of this item is finished. The case $\lambda < M$ is treated in the same way.

Corollary 5.4 *The only self-adjoint operator with a spectrum reduced to* $\{0\}$ *is the trivial operator* $\mathbf{T} \equiv \mathbf{0}$.

Proof By Theorem 5.5, if $\sigma(\mathbf{T}) = \{0\}$, we can conclude that the two numbers $m = M = 0$ vanish. Hence

$$\langle \mathbf{T}\mathbf{z}, \mathbf{z} \rangle = 0 \text{ for every } \mathbf{z} \in \mathbb{H}.$$

In particular, for arbitrary $\mathbf{x}$, $\mathbf{y}$,

$$\begin{aligned}
0 =& \langle \mathbf{T}(\mathbf{x} + \mathbf{y}), \mathbf{x} + \mathbf{y} \rangle \\
=& \langle \mathbf{T}\mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{T}\mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{T}\mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{T}\mathbf{y}, \mathbf{y} \rangle \\
=& 2\langle \mathbf{T}\mathbf{x}, \mathbf{y} \rangle,
\end{aligned}$$

and $\mathbf{T} \equiv \mathbf{0}$. □

Remark 5.2 In the context of the two previous results, it can be shown that for a self-adjoint operator $\mathbf{T}$, it is true that

$$\|\mathbf{T}\| = \max\{|m|, |M|\}.$$

5.7 The Fourier Transform

The utmost importance of the Fourier transform in Analysis can hardly be suitably estimated. In this section, we look at it from the viewpoint of the theory of continuous operators. It is just but a timid introduction to this fundamental integral transform.

Definition 5.7 For a function $f(\mathbf{x}) \in L^1(\mathbb{R}^N; \mathbb{C})$, we define

$$\mathcal{F}f(\mathbf{y}) = \hat{f}(\mathbf{y}) : \mathbb{R}^N \to \mathbb{C}$$

through

$$\hat{f}(\mathbf{y}) = \int_{\mathbb{R}^N} f(\mathbf{x}) e^{-2\pi i \mathbf{x}\cdot\mathbf{y}}\, d\mathbf{x}, \quad \mathbf{y} \in \mathbb{R}^N.$$

i is, of course, the imaginary unit. Recall that the inner product in $L^2(\mathbb{R}^N; \mathbb{C})$ is given by

$$\langle f, g\rangle = \int_{\mathbb{R}^N} f(\mathbf{x})\overline{g}(\mathbf{x})\, d\mathbf{x}$$

where $\bar{\cdot}$ represent the complex conjugate.

We immediately notice that for every individual $\mathbf{y} \in \mathbb{R}^N$, we have

$$|\hat{f}(\mathbf{y})| \le \int_{\mathbb{R}^N} |f(\mathbf{x})|\, d\mathbf{x},$$

and so

$$\|\hat{f}\|_{L^\infty(\mathbb{R}^N;\mathbb{C})} \le \|f\|_{L^1(\mathbb{R}^N;\mathbb{C})}.$$

However, the full power of the Fourier transform is shown when it is understood as an isometry in $L^2(\mathbb{R}^N; \mathbb{C})$. To this end, we state several basic tools, whose proof is deferred until we see how to apply them to our purposes, and are sure they deserve the effort.

Lemma 5.2 (Riemann-Lebesgue) *If* $f \in L^1(\mathbb{R}^N; \mathbb{C})$, *then*

$$\lim_{|\mathbf{y}|\to\infty} \int_{\mathbb{R}^N} f(\mathbf{x}) e^{-2\pi i \mathbf{x}\cdot\mathbf{y}}\, d\mathbf{x} = 0.$$

Lemma 5.3 (The Inverse Transform) *If* $f, \hat{f} \in L^1(\mathbb{R}^N; \mathbb{C})$, *and*

$$\lim_{r\to\infty} f(\mathbf{x}/r) = f(\mathbf{0}),$$

then

$$f(\mathbf{0}) = \int_{\mathbb{R}^N} \hat{f}(\mathbf{y})\, d\mathbf{y}.$$

In general, for almost every individual $\mathbf{x} \in \mathbb{R}^N$,

$$f(\mathbf{x}) = \int_{\mathbb{R}^N \times \mathbb{R}^N} f(z) e^{2\pi i (\mathbf{x}-z)\cdot \mathbf{y}}\, dz\, d\mathbf{y}.$$

This lemma permits the following coherent definition.

Definition 5.8 For a function $f \in L^1(\mathbb{R}^N; \mathbb{C})$, we define

$$\overline{\mathcal{F}}(f)(\mathbf{y}) \equiv \check{f}(\mathbf{y}) = \int_{\mathbb{R}^N} f(\mathbf{x}) e^{2\pi i \mathbf{x}\cdot\mathbf{y}}\, d\mathbf{x}.$$

The previous lemma can be interpreted in the sense that

$$\overline{\mathcal{F}}(f) = \mathcal{F}^{-1}(f),$$

when both transforms are appropriately interpreted in $L^2(\mathbb{R}^N; \mathbb{C})$.

Proposition 5.5 (Plancherel's Theorem in $L^1 \cap L^2$) *If $f, g, \hat{f}, \hat{g}$ belong to $L^1(\mathbb{R}^N; \mathbb{C}) \cap L^2(\mathbb{R}^N; \mathbb{C})$, then*

$$\langle \hat{f}, \hat{g} \rangle = \langle f, g \rangle.$$

In particular, $\|\hat{f}\|^2 = \|f\|^2$.

Define, for $f \in L^2(\mathbb{R}^N; \mathbb{C})$,

$$\mathcal{F}_j f(\mathbf{y}) = \int_{[-j,j]^N} f(\mathbf{x}) e^{-2\pi i \mathbf{x}\cdot\mathbf{y}}\, d\mathbf{x} = \mathcal{F}(f\chi_j),$$

where χ_j stands for the characteristic function of the box $[-j, j]^N$. This definition is legitimate because

$$f\chi_j \in L^1(\mathbb{R}^N; \mathbb{C}) \cap L^2(\mathbb{R}^N; \mathbb{C}).$$

Moreover, we claim that

$$\mathcal{F}(f\chi_j) \in L^1(\mathbb{R}^N; \mathbb{C}).$$

To show this, we use the Riemann-Lebesgue lemma for $\mathcal{F}(f\chi_j)(\mathbf{y})$ and all of its partial derivatives of any order. Indeed, if j is fixed, since $f\chi_j$ has compact support,

so does the product of it with every complex polynomial $P(\boldsymbol{x})$, and hence

$$f\chi_j P \in L^1(\mathbb{R}^N; \mathbb{C}).$$

In this case, we can differentiate $\mathcal{F}(f\chi_j)(\boldsymbol{y})$ with respect to $\boldsymbol{y}$ any number of times. For instance,

$$\frac{\partial \mathcal{F}(f\chi_j)}{\partial y_k} = \int_{\mathbb{R}^N} f(\boldsymbol{y})\chi_j(\boldsymbol{y})(-2\pi i x_k)e^{-2\pi i \boldsymbol{x}\cdot\boldsymbol{y}}\, d\boldsymbol{y}, \quad \boldsymbol{x} = (x_k)_k, \boldsymbol{y} = (y_k)_k.$$

This implies that $\mathcal{F}(f\chi_j)(\boldsymbol{y})$ is C^∞, and, thanks to Lemma 5.2 applied to $f\chi_j P$, the limit of all partial derivatives of $\mathcal{F}(f\chi_j)(\boldsymbol{y})$ vanishes as $|\boldsymbol{y}| \to \infty$. This, in particular, means that

$$\mathcal{F}(f\chi_j)(\boldsymbol{y}) \in L^1(\mathbb{R}^N; \mathbb{C}) \cap L^2(\mathbb{R}^N; \mathbb{C}).$$

Proposition 5.3 permits us to ensure that

$$\|\mathcal{F}(f\chi_j)\|^2 = \|f\chi_j\|^2 \le \|f\|^2 \tag{5.13}$$

for all j. This fact means that operators $\mathcal{F}_j$ are well-defined from $L^2(\mathbb{R}^N; \mathbb{C})$ to itself.

We are now ready to show the main fact of the theory of the Fourier transform in $L^2(\mathbb{R}^N; \mathbb{C})$. We will apply Corollaries 5.1 and 5.2.

Proposition 5.6 *The limit*

$$\mathcal{F} = \lim_{j\to\infty} \mathcal{F}_j \text{ in } L^2(\mathbb{R}^N; \mathbb{C})$$

defines an isometry in this same Hilbert space

$$\|\mathcal{F}f\| = \|f\|, \quad f \in L^2(\mathbb{R}^N; \mathbb{C}). \tag{5.14}$$

Identity (5.14) is known universally as Plancherel's identity.

Proof According to Corollary 5.1, we need to check that the limit of $\mathcal{F}_j f$ exists in $L^2(\mathbb{R}^N; \mathbb{C})$ for every individual function f in the same space. Let $j < k$ be two large positive integers. By our calculations above redone for the difference $f(\chi_j - \chi_k)$, we can conclude that

$$\|\mathcal{F}_j f - \mathcal{F}_k f\| = \|f(\chi_j - \chi_k)\| \to 0, \quad j, k \to \infty,$$

i.e. $\{\mathcal{F}_j f\}$ is a Cauchy sequence in $L^2(\mathbb{R}^N;\mathbb{C})$, and, as such, it converges to an element

$$\mathcal{F}f \in L^2(\mathbb{R}^N;\mathbb{C}).$$

By Corollary 5.2, $\mathcal{F}$ is a continuous operator in such Hilbert space, and, according to (5.13),

$$\|\mathcal{F}\| \le 1.$$

More is true because (5.13) implies that

$$\|\mathcal{F}(f\chi_{\mathcal{K}})\| = \|f\chi_{\mathcal{K}}\|$$

for every compact subset $\mathcal{K} \subset \mathbb{R}^N$, and, by density through the theorem of dominated convergence,

$$\|\mathcal{F}f\| = \|f\|.$$

□

We therefore see that the proofs of Lemmas 5.2, 5.3, and Proposition 5.14 are important. We now examine them.

***Proof** (Lemma 5.2)* This important fact is essentially Exercise 16 in Chap. 2. We argue in several steps:

1. Suppose first that f is the characteristic function of a box $\mathbf{B} = \Pi_k[a_k, b_k]$. Then

$$\begin{aligned}\int_{\mathbf{B}} e^{-2\pi i \boldsymbol{x}\cdot\boldsymbol{y}}\, d\boldsymbol{x} =& \Pi_k \int_{a_k}^{b_k} e^{-2\pi i x_k y_k}\, dx_k\\ =& \Pi_k \frac{1}{-2\pi i y_k}\, e^{-2\pi i x_k y_k}\Big|_{a_k}^{b_k} \to 0\end{aligned}$$

   as $y_k \to \infty$. Since at least one of the coordinates y_k must tend to infinity, the claim is correct.
2. By linearity, the conclusion is valid for linear combinations of characteristic functions of boxes.
3. By density, our claim becomes true for arbitrary functions in the space $L^1(\mathbb{R}^N;\mathbb{C})$. The argument goes like this. Take

$$f \in L^1(\mathbb{R}^N;\mathbb{C}),$$

and for arbitrary $\epsilon > 0$ find a linear combination χ of characteristic functions of boxes with $\|f - \chi\| \leq \epsilon$. In this manner

$$\int_{\mathbb{R}^N} f(\boldsymbol{x}) e^{-2\pi i \boldsymbol{x}\cdot\boldsymbol{y}}\, d\boldsymbol{x} = \int_{\mathbb{R}^N} (f(\boldsymbol{x}) - \chi(\boldsymbol{x})) e^{-2\pi i \boldsymbol{x}\cdot\boldsymbol{y}}\, d\boldsymbol{x} + \int_{\mathbb{R}^N} \chi(\boldsymbol{x}) e^{-2\pi i \boldsymbol{x}\cdot\boldsymbol{y}}\, d\boldsymbol{x}.$$

The first term cannot be any greater than ϵ, while the second can be made arbitrarily small by taking $\boldsymbol{y}$ sufficiently large, according to our previous step. □

***Proof** (**Lemma 5.3**)* We start looking at the Fourier transform of the special function

$$\phi_N(\boldsymbol{x}) = e^{-\pi|\boldsymbol{x}|^2} = \Pi_k e^{-i\pi x_k^2},$$

where its product structure enables us to focus on the one dimensional version

$$\phi_1(x) \equiv \phi(x) = e^{-\pi x^2}.$$

Once we have its Fourier transform $\hat{\phi}(y)$, then

$$\hat{\phi}_N(\boldsymbol{y}) = \Pi_k \hat{\phi}(y_k).$$

For the one-dimensional case, we write

$$\hat{\phi}(y) = \int_{\mathbb{R}} \phi(x) e^{-2\pi i x y}\, dx.$$

If we differentiate formally with respect to y, we are led to

$$\hat{\phi}'(y) = \int_{\mathbb{R}} \phi(x)(-2\pi i x) e^{-2\pi i x y}\, dy = i \int_{\mathbb{R}} \phi(x)(-2\pi x) e^{-2\pi i x y}\, dx.$$

If we notice that $\phi'(x) = -2\pi x \phi(x)$, then

$$\hat{\phi}'(y) = i \int_{\mathbb{R}} \phi'(x) e^{-2\pi i x y}\, dx,$$

and an integration by parts takes us to (contributions coming from infinity vanish)

$$\hat{\phi}'(y) = -2\pi y \int_{\mathbb{R}} \phi(x) e^{-2\pi i x y}\, dx = -2\pi y \hat{\phi}(y).$$

We conclude that $\hat{\phi}(y) = c\phi(y)$ for some constant c. If we examine this identity at $y = 0$, we see that

$$c = \hat{\phi}(0) = \int_{\mathbb{R}} \phi(x)\, dx = 1.$$

This last integral is well-known from elementary Calculus courses. Hence

$$\hat{\phi}(y) = \phi(y) = e^{-\pi y^2}, \quad \hat{\phi}_N(\boldsymbol{y}) = e^{-\pi |\boldsymbol{y}|^2}.$$

For positive r, if we put

$$g(\boldsymbol{x}) = \phi_N(\boldsymbol{x}/r),$$

we can easily check that

$$\int_{\mathbb{R}^N} f(\boldsymbol{x}/r)\hat{\phi}_N(\boldsymbol{x})\, d\boldsymbol{x} = \int_{\mathbb{R}^N} f(\boldsymbol{x})\hat{g}(\boldsymbol{x})\, d\boldsymbol{x}$$

by manipulating the change of scale in the natural manner. It is also elementary to realize, through Fubini's theorem, that the Fourier transform is formally self-adjoint in the sense

$$\int_{\mathbb{R}^N} \hat{f}(\boldsymbol{x})g(\boldsymbol{x})\, d\boldsymbol{x} = \int_{\mathbb{R}^N} f(\boldsymbol{x})\hat{g}(\boldsymbol{x})\, d\boldsymbol{x}, \mathit{and}$$

hence we arrive at

$$\int_{\mathbb{R}^N} f(\boldsymbol{x}/r)\hat{\phi}_N(\boldsymbol{x})\, d\boldsymbol{x} = \int_{\mathbb{R}^N} \hat{f}(\boldsymbol{x})\phi_N(\boldsymbol{x}/r)\, d\boldsymbol{x},$$

for every positive r. If we take r to infinity, the right-hand side converges, because $\hat{f} \in L^1(\mathbb{R}^N; \mathbb{C})$, to the integral of $\hat{f}$, and we obtain our claim.

For a.e. $\boldsymbol{x}$, we can apply the previous result translating the origen to the point $\boldsymbol{x}$, and so

$$f(\boldsymbol{x}) = \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} f(\boldsymbol{z} + \boldsymbol{x})e^{-2\pi i \boldsymbol{z}\cdot\boldsymbol{y}}\, d\boldsymbol{z}\, d\boldsymbol{y}.$$

A natural change of variables in the inner integral leads directly to our formula. □

Proof ***(Proposition 5.14)*** Plancherel's identity is a direct consequence of the more general fact concerning the behavior of Fourier's transform with respect to the inner product in $L^2(\mathbb{R}^N; \mathbb{C})$, namely, if

$$f, g \in L^2(\mathbb{R}^N; \mathbb{C}),$$

then

$$\langle \hat{f}, \hat{g} \rangle = \langle f, g \rangle.$$

In fact, the left-hand side is

$$\langle \hat{f}, \hat{g} \rangle = \int_{\mathbb{R}^N} \left[\int_{\mathbb{R}^N} f(\boldsymbol{x}) e^{-2\pi i \boldsymbol{x} \cdot \boldsymbol{y}} \, d\boldsymbol{x} \int_{\mathbb{R}^N} \overline{g}(\boldsymbol{z}) e^{2\pi i \boldsymbol{z} \cdot \boldsymbol{y}} \, d\boldsymbol{z} \right] d\boldsymbol{y},$$

and elementary manipulations lead to

$$\begin{aligned} \langle \hat{f}, \hat{g} \rangle &= \int_{\mathbb{R}^N} \left[\int_{\mathbb{R}^N \times \mathbb{R}^N} f(\boldsymbol{x}) \overline{g}(\boldsymbol{z}) e^{2\pi i (\boldsymbol{z} - \boldsymbol{x}) \cdot \boldsymbol{y}} \, d\boldsymbol{x} \, d\boldsymbol{z} \right] d\boldsymbol{y} \\ &= \int_{\mathbb{R}^N} \hat{g}(\boldsymbol{z}) \left[\int_{\mathbb{R}^N \times \mathbb{R}^N} f(\boldsymbol{x}) e^{2\pi i (\boldsymbol{z} - \boldsymbol{x}) \cdot \boldsymbol{y}} \, d\boldsymbol{y} \, d\boldsymbol{x} \right] d\boldsymbol{z}. \end{aligned}$$

We conclude directly by Lemma 5.3. □

In the proof of these fundamental facts we have used along the proof some of the basic properties that make the Fourier transform so useful in Applied Analysis: the behavior with respect to translations and changes of scale, the derivative of the transform, the transform of the derivative, etc.

As remarked at the beginning of the section, the importante of the Fourier transform goes well beyond what has been described here.

5.8 Exercises

1. Show that the two definitions of a Baire space in Definition 5.2 are indeed equivalent.
2. Prove that if

$$\mathbf{B}_\rho \subset \mathbf{T}(\mathbf{B}_1)$$

   for a linear, continuous map $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ between Banach spaces, and for some $\rho > 0$, then the image under $\mathbf{T}$ of every open subset of $\mathbb{E}$ is open in $\mathbb{F}$.
3. Show that if the graph of a linear operator $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ is closed in the product space, then $\mathbb{E}$ under the graph norm

$$\|\mathbf{u}\|_G = \|\mathbf{u}\|_{\mathbb{E}} + \|\mathbf{T}\mathbf{u}\|_{\mathbb{F}},$$

   is a Banach space too.

4. Let $\mathbb{E}$ be a Banach space, and $\mathbf{T} \in \mathcal{L}(\mathbb{E})$ such that $\|\mathbf{T}\| < 1$. Argue that $\mathbf{1} - \mathbf{T}$ is an isomorphism, by checking that the series
$$\mathbf{S} = \sum_{i=0}^{\infty} \mathbf{T}^i$$
is its inverse.
5. Consider the operator
$$\mathbf{T} : L^2(-\pi, \pi) \mapsto L^2(-\pi, \pi), \quad \mathbf{T}f(x) = f(2x),$$
where functions in $L^2(-\pi, \pi)$ are regarded as extended to all of $\mathbb{R}$ by periodicity.
   (a) Check that $\mathbf{T}$ is linear, continuous, and compute its norm.
   (b) Show that $0 \in \sigma(\mathbf{T}) \setminus e(\mathbf{T})$.
   (c) Calculate its adjoint $\mathbf{T}'$.
6. Let $\mathbb{H}$ be a Hilbert space.
   (a) If $\pi : \mathbb{H} \to \mathbb{H}$ is such that
$$\pi^2 = \pi, \quad \langle \pi \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \pi \mathbf{y} \rangle,$$
for every $\mathbf{x}, \mathbf{y} \in \mathbb{H}$, prove that π is linear continuous, and has norm 1. Conclude that $\mathbb{K} = \pi(\mathbb{H})$ is a closed subspace, orthogonal to the kernel of π, and π is the orthogonal projection onto $\mathbb{K}$.
   (b) If $\pi \in \mathcal{L}(\mathbb{H})$ is such that $\pi^2 = \pi$, argue that the four following assertions are equivalent: π is the orthogonal projection onto $\pi(\mathbb{H})$; π is positive; π is self-adjoint; and π commutes with π'.
7. For the operator
$$\mathbf{T} : \ell^p \to \ell^p, \quad \mathbf{T}(x_1, x_2, \dots) = (0, x_1, x_2, \dots),$$
show that
$$\sigma(\mathbf{T}) = \{|\lambda| \le 1\}, \quad e(\mathbf{T}) = \emptyset.$$
8. For the operator
$$\mathbf{T} : \ell^p \to \ell^p, \quad \mathbf{T}(x_1, x_2, x_3, \dots) = (x_1, x_2/2, x_3/3, \dots),$$
find $\sigma(\mathbf{T})$ and $e(\mathbf{T})$. For $\lambda \in \rho(\mathbf{T})$, calculate explicitly $(\mathbf{T} - \lambda \mathbf{1})^{-1}$.

9. Suppose

$$K(x, y) : [0, 1] \times [0, 1] \to \mathbb{R}, \quad K(x, y) \equiv 0 \text{ if } x < y,$$

is bounded, and continuous when $x \geq y$. The integral operator

$$\mathbf{T}u(x) = \int_0^x K(x, y)u(y)\,dy$$

is known as the Volterra operator with kernel K.

(a) Show that $\mathbf{T}$ is linear and continuous from $C[0, 1]$ to itself, and that the successive powers $\mathbf{T}^j$ also are Volterra operators.
(b) Show, by induction, that if K_j is the kernel of $\mathbf{T}^j$, then

$$|K_j(x, y)| \leq \frac{M^j}{(j-1)!}(x - y)^{j-1}$$

if $M > 0$ is an upper bound of $|K(x, y)|$.
(c) Look for examples of kernels $K(x, y)$ for which $0 \in e(\mathbf{T})$ and $0 \notin e(\mathbf{T})$.
(d) For the particular example

$$K(x, y) = \chi_\Delta(x, y), \quad \Delta = \{(x, y) \in [0, 1]^2 : x \geq y\},$$

the indicator function of Δ, calculate explicitly K_j, $\mathbf{T}^j$, and $(\mathbf{1} - \mathbf{T})^{-1}$. In this way, the equation $(\mathbf{1} - \mathbf{T})u = v$ can be resolved for $v \in C[0, 1]$.
(e) If v is differentiable, check that the equation $(\mathbf{1} - \mathbf{T})u = v$ is equivalent to a linear differential equation with a suitable initial condition, and that the corresponding solution in the previous item coincide with the one coming from the theory of Differential Equations.

10. Let $\mathbf{T} = \mathcal{L}(\mathbb{E})$, for a Banach space $\mathbb{E}$, such that the series

$$\sum_{j=0}^{\infty} \mathbf{T}^j(\mathbf{x})$$

converges in $\mathbb{E}$ for every $\mathbf{x} \in \mathbb{E}$ (note the difference with Exercise 4 above).

(a) Show that $\mathbf{1} - \mathbf{T}$ is an isomorphism of $\mathbb{E}$ and

$$(\mathbf{1} - \mathbf{T})^{-1}(\mathbf{y}) = \sum_{j=0}^{\infty} \mathbf{T}^j(\mathbf{y}).$$

(b) For arbitrary $\mathbf{x}_0 \in \mathbb{E}$, consider the sequence

$$\mathbf{x}_{j+1} = \mathbf{y} + \mathbf{T}\mathbf{x}_j.$$

Argue that it converges to $(\mathbf{1} - \mathbf{T})^{-1}(\mathbf{y})$, regardless of how $\mathbf{x}_0$ is selected.

(c) If we put

$$\mathbf{x} = \lim_{j \to \infty} \mathbf{x}_j, \quad \mathbf{x}_{j+1} = \mathbf{y} + \mathbf{T}\mathbf{x}_j,$$

show that, for all j,

$$\|\mathbf{x} - \mathbf{x}_j\| \leq \|(\mathbf{1} - \mathbf{T})^{-1}\| \, \|\mathbf{T}^j\| \, \|\mathbf{x}_1 - \mathbf{x}_0\|.$$

11. In the Banach space $\mathbb{E} = C[0, 1]$, endowed with the sup norm, let $\mathbb{K}$ be a subspace of differentiable functions (C^1). The derivative operator taking each $u \in \mathbb{K}$ into its derivative $u' \in \mathbb{E}$ is continuous if and only if $\mathbb{K}$ is finite-dimensional. In particular, if $\mathbb{K} \subset C^1[0, 1]$ is closed in $\mathbb{E}$, then it has finite dimension.
12. For functions

$$a(\boldsymbol{x}) \in L^\infty(\Omega), \quad b(\boldsymbol{x}) \in L^2(\Omega), \quad \Omega \subset \mathbb{R}^N,$$

define the operator $\mathbf{T} \in \mathcal{L}(L^2(\Omega))$ by putting

$$\mathbf{T}(f)(\boldsymbol{x}) = a(\boldsymbol{x}) f(\boldsymbol{x}) + b(\boldsymbol{x}).$$

13. The Laplace transform is universally defined through the formula

$$\mathbb{L}(u)(s) = \int_0^\infty u(t) e^{-st} \, dt, \quad s \in (0, \infty).$$

14. Use the Riesz representation theorem, Proposition 2.14 and the duality relations 5.3, to prove the non-symmetric version of the Lax-Milgram lemma: Let $A(\mathbf{u}, \mathbf{v})$ be a continuous, coercive bilinear form over a Hilbert space $\mathbb{H}$. For every $\mathbf{U} \in \mathbb{H}$, there is a unique $\overline{\mathbf{u}} \in \mathbb{H}$ such that

$$A(\overline{\mathbf{u}}, \mathbf{v}) = \langle \mathbf{U}, \mathbf{v} \rangle.$$

for every $\mathbf{v} \in \mathbb{H}$.
15. For $\mathbb{E} = C[0, 1]$ under the sup-norm, consider the operator $\mathbf{T}$ taking each $y(t) \in \mathbb{E}$ into the unique solution $x(t)$ of the differential problem

$$x'(t) = x(t) + y(t) \text{ in } [0, 1], \quad x(0) = 0.$$

Check that it is linear, continuous, and find its norm, and its eigenvalues.

Chapter 6
Compact Operators

Compactness of an operator goes one step beyond continuity. Its treatment is fundamental in many areas in Applied Analysis, particularly Differential Equations and Variational Methods. Since compactness is one of the main concepts in Topology when applied to Analysis, compact operators will also play a fundamental role. In some sense, the class of compact operators is the one that allows to shift our intuition about finite-dimensional spaces to the infinite-dimensional scenario.

6.1 Preliminaries

We start with the basic definition, properties, and examples.

Definition 6.1 A linear, continuous operator $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ between two Banach spaces is said to be compact if the image of the unit ball $\mathbf{B}_{\mathbb{E}}$ of $\mathbb{E}$ is relatively compact in $\mathbb{F}$. Equivalently, if from every uniformly bounded sequence $\{\mathbf{u}_j\}$ in $\mathbb{E}$, one can find a subsequence $\{\mathbf{T}\mathbf{u}_{j_k}\}$ converging in $\mathbb{F}$. The set of compact operators from $\mathbb{E}$ to $\mathbb{F}$ is denoted by $\mathcal{K}(\mathbb{E}, \mathbb{F})$.

Because uniformly bounded sequences in reflexive Banach spaces admit subsequences converging weakly, roughly speaking compact operators are those transforming weak into strong convergence.

Proposition 6.1 *For two given Banach spaces $\mathbb{E}$ and $\mathbb{F}$, $\mathcal{K}(\mathbb{E}, \mathbb{F})$ is a closed subspace of $\mathcal{L}(\mathbb{E}, \mathbb{F})$.*

Proof The fact that $\mathcal{K}(\mathbb{E}, \mathbb{F})$ is a subspace of $\mathcal{L}(\mathbb{E}, \mathbb{F})$ is pretty obvious. Let

$$\{\mathbf{T}_k\} \subset \mathcal{K}(\mathbb{E}, \mathbb{F}), \quad \mathbf{T}_k \to \mathbf{T} \text{ in } \mathcal{L}(\mathbb{E}, \mathbb{F}),$$

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4_6

and $\{\mathbf{u}_j\} \subset \mathbb{E}$ uniformly bounded. Since, $\mathbb{F}$ is a complete metric space, it is a well-known fact in Topology, that it suffices to check that the image $\mathbf{T}(\mathbf{B}_{\mathbb{E}})$ can be covered by a finite number of balls of arbitrarily small size. Take $\epsilon > 0$, and k with

$$\|\mathbf{T}_k - \mathbf{T}\| < \epsilon/2.$$

Since $\mathbf{T}_k$ is compact, we can find $\mathbf{v}_i \in \mathbb{F}$ with

$$\mathbf{T}_k(\mathbf{B}_{\mathbb{E}}) \subset \cup_i \mathbf{B}(\mathbf{v}_i, \epsilon/2).$$

The triangular inequality immediately leads to

$$\mathbf{T}(\mathbf{B}_{\mathbb{E}}) \subset \cup_i \mathbf{B}(\mathbf{v}_i, \epsilon),$$

and the arbitrariness of ϵ implies $\mathbf{T} \in \mathcal{K}(\mathbb{E}, \mathbb{F})$. □

There are several main initial examples of compact operators.

Example 6.1 A linear operator $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ is of finite rank if the image subspace $\mathbf{T}(\mathbb{E})$ is finitely generated. Every linear operator of finite rank is compact. According to the previous proposition, every limit operator of a sequence of finite rank operators is also compact. The converse of this statement is known as the approximation problem: to decide whether every compact operator is a limit of a sequence of finite-rank operators. Though the answer is positive when, for instance, $\mathbb{F}$ is a Hilbert space, it is indeed a very delicate issue which is not true in general.

Example 6.2 Let $J = (x_0, x_1)$ be a finite interval. The integration operator

$$\mathcal{I} : W^{1,p}(J; \mathbb{R}^n) \to L^\infty(J; \mathbb{R}^n), \quad \mathcal{I}\mathbf{u}(x) = \int_{x_0}^{x} \mathbf{u}'(y)\,dy,$$

for $p > 1$ is compact. We already know this from Chap. 2.

Example 6.3 The prototype of an operation which is compact is integration, as opposed to differentiation, as we have started to suspect from the previous example. To clearly show this assertion, let $\Omega \subset \mathbb{R}^N$ be a bounded open set, and let

$$K(\mathbf{x}, \mathbf{y}) : \Omega \times \Omega \to \mathbb{R}, \quad |K(\mathbf{x}, \mathbf{y}) - K(\mathbf{z}, \mathbf{y})| \le M|\mathbf{x} - \mathbf{z}|^\alpha k(\mathbf{y}),$$
$$k(\mathbf{y}) \in L^2(\Omega),$$

for a positive constant M, and exponent $\alpha > 0$. Define the operator

$$\mathbf{T} : L^2(\Omega) \to L^\infty(\Omega), \quad \mathbf{T}u(\mathbf{x}) = \int_\Omega K(\mathbf{x}, \mathbf{y})u(\mathbf{y})\,d\mathbf{y}.$$

$\mathbf{T}$ is compact. This is easy to see because for a bounded sequence $\{u_j\}$ in $L^2(\Omega)$, we can write for an arbitrary pair of points $\mathbf{x}$ and $\mathbf{z}$,

$$|\mathbf{T}u_j(\mathbf{x}) - \mathbf{T}u_j(\mathbf{z})| \le \tilde{M}|\mathbf{x} - \mathbf{z}|^\alpha \|k\|_{L^2(\Omega)}$$

where constant $\tilde{M}$ incorporates constant M and the uniform bound for the sequence $\{u_j\}$. This final inequality shows that the sequence of images $\{\mathbf{T}u_j\}$ is equicontinuous, and hence it admits some convergent subsequence.

Example 6.4 The compactness of an operator may depend on the spaces where it is considered. A convolution operator of the kind

$$\mathbf{T}u(\mathbf{x}) = \int_{\mathbb{R}^N} \rho(\mathbf{x} - \mathbf{y})u(\mathbf{y})\, d\mathbf{y}$$

with a smooth, compactly supported kernel

$$\rho(\mathbf{z}) : \mathbb{R}^N \to \mathbb{R},$$

is compact when regarded as

$$\mathbf{T} : L^1(\mathbb{R}^N) \to L^\infty(\mathbb{R}^N).$$

It is a particular example of the situation in the preceding example for a kernel

$$K(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{x} - \mathbf{y}).$$

However, it is not compact when considered from $L^1(\mathbb{R}^N)$ to itself, because the injection

$$L^\infty(\mathbb{R}^N) \cap L^1(\mathbb{R}^N) \subset L^1(\mathbb{R}^N)$$

is not compact. If we restrict attention to a subspace $L^1(\Omega)$ for a bounded subset $\Omega \subset \mathbb{R}^N$, then it is compact though.

Example 6.5 The kind of integral operators considered in the last examples can be viewed in a much more general framework, and still retain its compactness feature. These are called Hilbert-Schmidt operators. Let

$$K(\mathbf{x}, \mathbf{y}) : \Omega \times \Omega \to \mathbb{R}, \quad \Omega \subset \mathbb{R}^N,$$

with

$$K(\mathbf{x}, \mathbf{y}) \in L^2(\Omega \times \Omega), \quad \int_{\Omega \times \Omega} |K(\mathbf{x}, \mathbf{y})|^2 \, d\mathbf{x}\, d\mathbf{y} < \infty,$$

and define the corresponding integral operator as before

$$\mathbf{T}u(x) = \int_{\Omega} K(\mathbf{x}, \mathbf{y})u(\mathbf{y})\,d\mathbf{y}, \quad \mathbf{T} : L^2(\Omega) \to L^2(\Omega).$$

It is easy to check that $\mathbf{T}$ is continuous. To show its compactness, we take an orthonormal basis $\{v_j\}$ for $L^2(\Omega)$, and realize that the set of all products $\{v_i(\mathbf{x})v_j(\mathbf{y})\}$ makes a orthonormal basis for $L^2(\Omega \times \Omega)$. In this way we can write

$$K(x, y) = \sum_{i,j} k_{i,j} v_i(\mathbf{x})v_j(\mathbf{y}), \quad \sum_{i,j} |k_{i,j}|^2 < \infty,$$

and

$$\mathbf{T}u(\mathbf{x}) = \sum_i \left(\sum_j k_{i,j} \int_{\Omega} v_j(\mathbf{y})u(\mathbf{y})\,d\mathbf{y} \right) v_i(\mathbf{x}).$$

For a finite k, define

$$\mathbf{T}_k : L^2(\Omega) \to L^2(\Omega), \quad \mathbf{T}_k u(\mathbf{x}) = \sum_i^k \left(\sum_j^k k_{i,j} \int_{\Omega} v_j(\mathbf{y})u(\mathbf{y})\,d\mathbf{y} \right) v_i(\mathbf{x}).$$

It is clear that $\mathbf{T}_k$ is a finite-rank operator; and, on the other hand, the condition on the summability of the double series for the coefficients $k_{i,j}$ (remainders tend to zero) directly implies that

$$\|\mathbf{T}_k - \mathbf{T}\| \to 0, \quad k \to \infty.$$

Conclude by Proposition 6.1.

Proposition 5.3 establishes a clear relationship between the kernel and range of an operator $\mathbf{T}$ and its adjoint $\mathbf{T}'$. In order to make use of such identities in the context of compact operators, it is important to know whether the adjoint of a compact operator is compact as well.

Theorem 6.1 *Let $\mathbf{T} \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ for a couple of Banach spaces $\mathbb{E}$ and $\mathbb{F}$. Then $\mathbf{T} \in \mathcal{K}(\mathbb{E}, \mathbb{F})$ if and only if $\mathbf{T}' \in \mathcal{K}(\mathbb{F}', \mathbb{E}')$.*

Proof Suppose $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ is compact. Let $\{\mathbf{y}_k\}$ be a sequence in the unit ball $\mathbb{B}'$ of $\mathbb{F}'$. We would like to show that $\{\mathbf{T}'(\mathbf{y}_k)\}$ admits some convergent subsequence. Our main assumption, the compactness of $\mathbf{T}$, implies that $\mathbb{K} = \mathbf{T}(\mathbb{B})$ is compact in $\mathbb{F}$, if $\mathbb{B}$ is the unit ball of $\mathbb{E}$. Each $\mathbf{y}_k$ is, in particular, continuous in the compact set $\mathbb{K}$, and because $\|\mathbf{y}_k\| \leq 1$ for all k ($\mathbf{y}_k \in \mathbb{B}'$), the whole family $\{\mathbf{y}_k\}$ is equicontinuous when restricted to $\mathbb{K}$. By the classical Arzelá-Ascoli theorem, there is a convergent

subsequence (not relabeled), and hence

$$\sup_{\mathbf{x}\in\mathbb{B}} |\langle \mathbf{y}_j, \mathbf{T}\mathbf{x}\rangle - \langle \mathbf{y}_k, \mathbf{T}\mathbf{x}\rangle| \to 0$$

as $k, j \to \infty$. This is exactly the same as saying

$$\|\mathbf{T}'\mathbf{y}_j - \mathbf{T}'\mathbf{y}_k\| \to 0$$

as $k, j \to \infty$; that is $\{\mathbf{T}'\mathbf{y}_k\}$ is convergent in $\mathbb{E}'$.

Assume now that $\mathbf{T}' : \mathbb{F}' \to \mathbb{E}'$ is compact. What we have just proved implies that $\mathbf{T}'' : \mathbb{E}'' \to \mathbb{F}''$ is compact. By our remarks at the end of Sect. 5.4, we know that $\mathbf{T}(\mathbb{B}) = \mathbf{T}''(\mathbb{B})$ if $\mathbb{B}$ is the unit ball of $\mathbb{E}$. Moreover, $\mathbf{T}''(\mathbb{B})$ has a compact closure in $\mathbb{F}''$, and $\mathbb{F} \subset \mathbb{F}''$ is a closed subspace. We can then deduce that $\mathbf{T}(\mathbb{B})$ has compact closure in $\mathbb{F}$, and $\mathbf{T} : \mathbb{E} \to \mathbb{F}$ is compact. □

6.2 The Fredholm Alternative

For a Banach space $\mathbb{E}$, we designate by $\mathcal{K}(\mathbb{E})$ the class of compact operators from $\mathbb{E}$ into itself. A first fundamental result with a couple of basic facts about compact operators, that is the basis for the Fredholm alternative, follows. Recall that $\mathbf{1}$ is the identity map.

Proposition 6.2 *Let* $\mathbf{T} \in \mathcal{K}(\mathbb{E})$.

1. *If* $\mathbf{T}(\mathbb{E})$ *is a closed subspace, then it has finite dimension.*
2. *The subspaces* $\mathbb{N}(\mathbf{T} - \mathbf{1})$ *and* $\mathbb{N}(\mathbf{T}' - \mathbf{1})$ *always have finite dimension (and it is the same for both).*

Proof If $\mathbf{T}(\mathbb{E})$ is closed, it is a Banach space itself, and $\mathbf{T} : \mathbb{E} \to \mathbf{T}(\mathbb{E})$ turns out to be onto. According to the open mapping theorem, $\mathbf{T}$ is open, and so $\mathbf{T}(\mathbb{B})$, the image under $\mathbf{T}$ of the unit ball $\mathbb{B}$ in $\mathbb{E}$, is a relatively compact, neighborhood of zero in the space $\mathbf{T}(\mathbb{E})$. In this way, there is a finite set $\{\mathbf{u}_i\} \subset \mathbf{T}(\mathbb{E})$ such that

$$\mathbf{T}(\mathbb{B}) \subset \bigcup_i \left(\mathbf{u}_i + \frac{1}{2}\mathbf{T}(\mathbb{B})\right) \subset \mathbb{M} + \frac{1}{2}\mathbf{T}(\mathbb{B})$$

if $\mathbb{M}$ is the finite-dimensional space spanned by $\{\mathbf{u}_i\}$. By Exercise 2 below, we conclude that $\mathbf{T}(\mathbb{E})$ has finite dimension.

The subspace $\mathbb{N}(\mathbf{T}-\mathbf{1})$ is closed, and the compact restriction of $\mathbf{T}$ to this subspace is the identity. Thus the image subspace $\mathbf{T}(\mathbb{N}) = \mathbb{N}$ is closed. By the previous item, such subspace must be finite dimensional. The same applies to the compact operator $\mathbf{T}'$.

□

We are ready to deal with one of the most remarkable facts about compact operators.

Theorem 6.2 (Fredholm Alternative) *Let* $\mathbf{T} \in \mathcal{K}(\mathbb{E})$*, and* $\lambda \in \mathbb{R} \setminus \{0\}$*. Then*

$$\mathbb{R}(\mathbf{T} - \lambda \mathbf{1}) = \mathbb{N}(\mathbf{T}' - \lambda \mathbf{1})^{\perp}.$$

The usual, and useful, interpretation of this statement refers to the solvability of the equation

$$\mathbf{T}\mathbf{x} - \lambda \mathbf{x} = \mathbf{y}, \quad \mathbf{y} \in \mathbb{E}, \lambda \neq 0, \tag{6.1}$$

in two “alternatives":

1. either there is a unique solution $\mathbf{x}$ for (6.1) for every $\mathbf{y} \in \mathbb{E}$, in case $\mathbb{N}(\mathbf{T}' - \mathbf{1})$ is trivial; or else
2. Equation (6.1) for $\mathbf{y} = \mathbf{0}$ has a finite number of independent solutions, and the non-homogeneous equation (for $\mathbf{y} \neq \mathbf{0}$) admits solutions if and only if

$$\mathbf{y} \in \mathbb{N}(\mathbf{T}' - \lambda \mathbf{1})^{\perp} : \quad \langle \mathbf{y}, \mathbf{x}' \rangle = 0 \text{ whenever } \mathbf{T}'\mathbf{x}' - \lambda \mathbf{x}' = \mathbf{0}.$$

Proof By Proposition 5.3, it suffices to show that $\mathbb{R}(\mathbf{T} - \lambda \mathbf{1})$ is closed. Since $\mathbf{T}$ is compact if and only if $(1/\lambda)\mathbf{T}$ is compact for non-zero λ, we can, without loss of generality, concentrate in showing that $\mathbb{R}(\mathbf{T} - \mathbf{1})$ is closed.

Assume then that

$$\mathbf{T}\mathbf{x}_j - \mathbf{x}_j \to \mathbf{X}, \quad \mathbf{X} \in \mathbb{E}. \tag{6.2}$$

We are supposed to show the existence of some $\mathbf{x} \in \mathbb{E}$ with

$$\mathbf{T}\mathbf{x} - \mathbf{x} = \mathbf{X}.$$

Note that for arbitrary

$$\mathbf{z}_j \in \mathbb{N}(\mathbf{T} - \mathbf{1}),$$

we can replace $\mathbf{x}_j$ by $\mathbf{x}_j + \mathbf{z}_j$ without changing the limit $\mathbf{X}$

$$\mathbf{T}(\mathbf{x}_j + \mathbf{z}_j) - (\mathbf{x}_j + \mathbf{z}_j) \to \mathbf{X}. \tag{6.3}$$

In addition, by Proposition 6.2, the subspace $\mathbb{N}(\mathbf{T} - \mathbf{1})$ is finite-dimensional.

- We claim that $\mathbf{z}_j$ in the kernel of $\mathbf{T} - \mathbf{1}$ can be chosen so that $\mathbf{x}_j + \mathbf{z}_j$ remains a bounded sequence. To this aim, consider the function

$$\|\mathbf{x}_j + \cdot\| : \mathbb{N}(\mathbf{T} - \mathbf{1}) \to \mathbb{R}^{+}.$$

This is a convex, coercive function defined over a finite dimensional space. It therefore attains its global minimum at some $\mathbf{z}_j \in \mathbb{N}(\mathbf{T}-\mathbf{1})$ (possibly non-unique if the norm is not strictly convex). Suppose the sequence of numbers $\{\alpha_j = \|\mathbf{x}_j + \mathbf{z}_j\|\}$ is not bounded. In this case,

$$\frac{1}{\alpha_j}\mathbf{T}(\mathbf{x}_j + \mathbf{z}_j) - \frac{1}{\alpha_j}(\mathbf{x}_j + \mathbf{z}_j) \to \mathbf{0}. \tag{6.4}$$

But if, after extracting a subsequence not relabeled, due again to the compactness of $\mathbf{T}$,

$$\frac{1}{\alpha_j}\mathbf{T}(\mathbf{x}_j + \mathbf{z}_j) \to \mathbf{z},$$

then from (6.4),

$$\frac{1}{\alpha_j}(\mathbf{x}_j + \mathbf{z}_j) \to \mathbf{z}$$

as well, so that $\mathbf{Tz} = \mathbf{z}$, and $\mathbf{z} \in \mathbb{N}(\mathbf{T}-\mathbf{1})$. But because $\mathbb{N}(\mathbf{T}-\mathbf{1})$ is a subspace and $\mathbf{z}_j$ belongs to it, on the one hand,

$$\begin{aligned} &\min_{\mathbf{u}\in\mathbb{N}(\mathbf{T}-\mathbf{1})} \|(1/\alpha_j)(\mathbf{x}_j + \mathbf{z}_j) + \mathbf{u}\| \\ &= (1/\alpha_j) \min_{\mathbf{u}\in\mathbb{N}(\mathbf{T}-\mathbf{1})} \|\mathbf{x}_j + \mathbf{u}\| = 1, \end{aligned} \tag{6.5}$$

because this last minimum was attained at $\mathbf{z}_j$; but on the other, by taking $\mathbf{u} = -\mathbf{z}$ on the minimum on the left-hand side, we see that such an minimum must be smaller than

$$\|(1/\alpha_j)(\mathbf{x}_j + \mathbf{z}_j) - \mathbf{z}\|$$

which tends to zero, a contradiction with (6.5). Hence, $\{\mathbf{x}_j + \mathbf{z}_j\}$ is bounded.

- By the compactness of $\mathbf{T}$, for a subsequence which we do not care to relabel,

$$\mathbf{T}(\mathbf{x}_j + \mathbf{z}_j) \to \mathbf{Y}, \quad \text{some } \mathbf{Y} \in \mathbb{E}.$$

By (6.3),

$$\mathbf{x}_j + \mathbf{z}_j \to \mathbf{Y} - \mathbf{X},$$

and, then,

$$\mathbf{T}(\mathbf{x}_j + \mathbf{z}_j) \to \mathbf{T}(\mathbf{Y} - \mathbf{X}).$$

Putting these last two convergences together, and bearing in mind (6.2),

$$\mathbf{T}(\mathbf{Y}-\mathbf{X})-(\mathbf{Y}-\mathbf{X})=\mathbf{X},$$

i.e. $\mathbf{X}\in\mathbb{R}(\mathbf{T}-\mathbf{1})$, as desired.

□

Remark 6.1 The fact that the dimension is the same for the two finite-dimensional spaces $\mathbb{N}(\mathbf{T}-\mathbf{1})$ and $\mathbb{N}(\mathbf{T}'-\mathbf{1})$, when $\mathbf{T}$ is compact, can be shown in a elementary way by using Theorem 6.2 (and also Proposition 5.3, and the fact that finite-dimensional subspaces always admit complement).

6.3 Spectral Analysis

Recall that (Definition 5.5) a real number λ is an eigenvalue of an operator $\mathbf{T}\in\mathcal{L}(\mathbb{E})$, and we write $\lambda\in e(\mathbf{T})$, if the subspace $\mathbb{N}(\mathbf{T}-\lambda\mathbf{1})$ is not the trivial subspace. The resolvent $\rho(\mathbf{T})$ is the set of numbers λ for which $\mathbf{T}-\lambda\mathbf{1}$ is a bijection. Finally, the spectrum $\sigma(\mathbf{T})$ is the complement of the resolvent.

Before we discuss the main result in this section, we prove a very helpful tool. The distance function to a set $\mathbb{U}$ in a Banach space $\mathbb{E}$ is taken to be

$$d(\mathbf{x},\mathbb{U})=\inf_{\mathbf{y}\in\mathbb{U}}\|\mathbf{x}-\mathbf{y}\|,\quad \mathbf{x}\in\mathbb{E}.$$

Lemma 6.1 (Riesz) *Let $\mathbb{E}$ be a Banach space, and $\mathbb{M}$, a proper, closed subspace. For every $\epsilon>0$, there is $\mathbf{x}\in\mathbb{E}$ with $\|\mathbf{x}\|=1$ and such that*

$$d(\mathbf{x},\mathbb{M})\geq 1-\epsilon.$$

Proof By Corollary 3.3 of the Hahn-Banach theorem, there is $T\in\mathbb{E}'$ with

$$\|T\|=1,\quad T|_{\mathbb{M}}=\mathbf{0}.$$

For $\epsilon>0$, there is some unit $\mathbf{x}$, $\|\mathbf{x}\|=1$, such that

$$\|T\|-\epsilon\leq|\langle T,\mathbf{x}\rangle|.$$

Then, for every $\mathbf{y}\in\mathbb{M}$,

$$\begin{aligned}1-\epsilon=&\|T\|-\epsilon\leq|\langle T,\mathbf{x}\rangle|\\ =&|\langle T,\mathbf{x}-\mathbf{y}\rangle|\leq\|T\|\,\|\mathbf{x}-\mathbf{y}\|=\|\mathbf{x}-\mathbf{y}\|.\end{aligned}$$

This is the conclusion of the lemma because $\mathbf{y}\in\mathbb{M}$ is arbitrary. □

This lemma will be an indispensable tool for yet an intermediate step for our main theorem which is a revision of some spectral concepts for compact operators. It very clearly expresses the remarkable and surprising consequences of compactness.

Proposition 6.3 *Let* $\mathbf{T} \in \mathcal{K}(\mathbb{E})$. $\mathbb{N}(\mathbf{T}-\mathbf{1})$ *is the trivial subspace if and only if* $\mathbb{R}(\mathbf{T}-\mathbf{1})$ *the full space, i.e.* $\mathbf{T}-\mathbf{1}$ *is injective if and only if it is surjective.*

Proof Suppose first that $\mathbb{N}(\mathbf{T}-\mathbf{1})$ is trivial but $\mathbb{R}(\mathbf{T}-\mathbf{1})$ is not the full $\mathbb{E}$. By Theorem 6.2, the proper subspace $\mathbb{M} \equiv \mathbb{R}(\mathbf{T}-\mathbf{1})$ is closed, and so $\mathbb{M}$ is a Banach space on its own. Moreover, it is evident that $\mathbf{T}(\mathbb{M}) \subset \mathbb{M}$, but $\mathbf{T}(\mathbb{M})$ cannot fill up all of $\mathbb{M}$, precisely because $\mathbb{N}(\mathbf{T}-\mathbf{1})$ is trivial and $\mathbb{R}(\mathbf{T}-\mathbf{1})$ is not the full $\mathbb{E}$ (this is exactly as in Linear Algebra in finite dimension). In addition,

$$\mathbf{T}|_{\mathbb{M}} \in \mathcal{K}(\mathbb{M}).$$

We therefore see that we can define recursively the following sequence of subspaces

$$\mathbb{M}_j = (\mathbf{T}-\mathbf{1})^j \mathbb{E}, \quad j \geq 0,$$

in such a way that it is a strictly nested, decreasing sequence of closed subspaces (for the same reason indicated above). By Lemma 6.1, we can find a sequence $\{\mathbf{x}_j\}$ in such a way that

$$\mathbf{x}_j \in \mathbb{M}_j, \quad \|\mathbf{x}_j\| = 1, \quad \operatorname{dist}(\mathbf{x}_j, \mathbb{M}_{j+1}) \geq \frac{1}{2}. \tag{6.6}$$

If for $j < k$, we look at the difference

$$\mathbf{T}\mathbf{x}_j - \mathbf{T}\mathbf{x}_k = (\mathbf{T}\mathbf{x}_j - \mathbf{x}_j) - (\mathbf{T}\mathbf{x}_k - \mathbf{x}_k) + (\mathbf{x}_j - \mathbf{x}_k),$$

we realize that

$$(\mathbf{T}\mathbf{x}_j - \mathbf{x}_j) - (\mathbf{T}\mathbf{x}_k - \mathbf{x}_k) - \mathbf{x}_k \in \mathbb{M}_{j+1},$$

and hence

$$\mathbf{T}\mathbf{x}_j - \mathbf{T}\mathbf{x}_k = \mathbf{x}_j + \mathbf{z}, \quad \mathbf{z} \in \mathbb{M}_{j+1}.$$

But (6.6) implies, then, that

$$\|\mathbf{T}\mathbf{x}_j - \mathbf{T}\mathbf{x}_k\| \geq \frac{1}{2}$$

for all j, k, which is impossible for a compact operator. This contradiction leads us to conclude that, indeed, $\mathbb{R}(\mathbf{T}-\mathbf{1})$ is the full space.

The converse can be argued through the adjoint $\mathbf{T}'$, the dual $\mathbb{E}'$ and the relations in Proposition 5.3. It is left as an exercise. □

We are now ready for our main result concerning important spectral facts for compact operators.

Theorem 6.3 *Let $\mathbb{E}$ be a (non-trivial) Banach space, and $\mathbf{T} \in \mathcal{K}(\mathbb{E})$.*

1. $0 \in \sigma(\mathbf{T})$*;*
2. *every non-vanishing λ in $\sigma(\mathbf{T})$ is an eigenvalue of $\mathbf{T}$:*

$$\sigma(\mathbf{T}) \setminus \{0\} \subset e(\mathbf{T}) \quad or \quad \sigma(\mathbf{T}) = e(\mathbf{T}) \cup \{0\};$$

3. *for positive $r > 0$, the set*

$$E_r = \{\lambda \in \sigma(\mathbf{T}) : |\lambda| \geq r\}$$

   is empty or finite;
4. *the spectrum $\sigma(\mathbf{T})$ is a non-empty, countable, compact set of numbers with 0 as the only possible accumulation point.*

Proof If $0 \in \rho(\mathbf{T})$, then $\mathbf{T}$ should be a bijection but this is impossible for a compact operator on an infinite-dimensional Banach space.

Suppose a non-vanishing λ is not an eigenvalue, so that $\mathbf{T} - \lambda\mathbf{1}$ is injective. By Proposition 6.3, it is also surjective, and so $\lambda \in \rho(\mathbf{T})$.

For the third item, we follow a proof similar to that of Proposition 6.3. Let, for some fixed, positive $r > 0$, the set E_r be infinite: $\{\lambda_i\} \subset E_r$. By the previous item, each λ_i is an eigenvalue of $\mathbf{T}$ with some eigenvector $\mathbf{e}_i$ in such a way that the set $\{\mathbf{e}_i\}$ is a linearly independent set. Put $\mathbb{M}_n$ for the finite-dimensional subspace spanned by $\{\mathbf{e}_i\}_{i=1,2,\ldots,n}$ so that

$$\mathbb{M}_1 \subseteq \mathbb{M}_2 \subseteq \cdots \subseteq \mathbb{M}_n \subseteq \ldots$$

with proper inclusions. By Lemma 6.1, we can find a sequence $\{\mathbf{x}_i\}$ with

$$\|\mathbf{x}_i\| = 1, \quad \mathbf{x}_i \in \mathbb{M}_i, \quad \|\mathbf{x}_i - \mathbf{y}\| \geq \frac{1}{2}$$

for every $\mathbf{y} \in \mathbb{M}_{i-1}$. We claim, then, that the sequence $\{\mathbf{T}\mathbf{x}_i\}$ cannot possess a convergent subsequence, which is impossible due to the compactness of $\mathbf{T}$. Indeed, we can write for $j < k$,

$$\|\mathbf{T}\mathbf{x}_j - \mathbf{T}\mathbf{x}_k\| = \|\mathbf{T}\mathbf{x}_j - (\mathbf{T} - \lambda_k\mathbf{1})\mathbf{x}_k - \lambda_k\mathbf{x}_k\|.$$

Because $(\mathbf{T} - \lambda_k \mathbf{1})\mathbf{x}_k$ cannot have a non-vanishing component along $\mathbf{e}_k$, precisely because $\mathbf{e}_k$ is an eigenvector associated with the eigenvalue λ_k, we see that $(j < k)$

$$\mathbf{T}\mathbf{x}_j - (\mathbf{T} - \lambda_k \mathbf{1})\mathbf{x}_k \in \mathbb{M}_{k-1}$$

and so does this same vector divided by λ_k. In this way

$$\|\frac{1}{\lambda_k}(\mathbf{T}\mathbf{x}_j - (\mathbf{T} - \lambda_k \mathbf{1})\mathbf{x}_k) - \mathbf{x}_k\| \geq \frac{1}{2},$$

or

$$\|\mathbf{T}\mathbf{x}_j - \mathbf{T}\mathbf{x}_k\| = \|\mathbf{T}\mathbf{x}_j - (\mathbf{T} - \lambda_k \mathbf{1})\mathbf{x}_k - \lambda_k \mathbf{x}_k\| \geq \frac{|\lambda_k|}{2} \geq \frac{r}{2}.$$

This is impossible for a compact operator $\mathbf{T}$, and this contradiction implies that each E_r must be finite (or empty).

The last item is a direct consequence of the previous ones. □

6.4 Spectral Decomposition of Compact, Self-Adjoint Operators

We would like to explore a main fact that is a consequence of putting together these two fundamental properties, self-adjointness and compactness, in a given operator $\mathbf{T}$ from a Hilbert space into itself. It is the most fundamental property that we would like to highlight: they are always diagonalizable.

Theorem 6.4 *Let* $\mathbf{T} \in \mathcal{K}(\mathbb{H})$ *be a compact, self-adjoint operator over a separable Hilbert space* $\mathbb{H}$. *There exists a orthonormal basis of* $\mathbb{H}$ *made up of eigenvectors of* $\mathbf{T}$.

Proof By Theorem 6.3, there is a sequence of non-zero eigenvalues λ_j converging to 0. By Proposition 6.2 each subspace $\mathbb{N}(\mathbf{T} - \lambda_j \mathbf{1})$ is of finite dimension. Put

$$\mathbb{H}_0 = \mathbb{N}(\mathbf{T}), \quad \mathbb{H}_j = \mathbb{N}(\mathbf{T} - \lambda_j \mathbf{1}),\ j \geq 1.$$

We will show that $\mathbb{H}$ is the closure of the union of all the $\mathbb{H}_j$'s, and they are mutually orthogonal. Indeed, if $\mathbf{x} \in \mathbb{H}_j$ and $\mathbf{y} \in \mathbb{H}_k$ with $j \neq k$, then

$$(\lambda_j - \lambda_k)\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{T}\mathbf{x}, \mathbf{y}\rangle - \langle \mathbf{x}, \mathbf{T}\mathbf{y}\rangle = 0$$

because $\mathbf{T}$ is self-adjoint, and

$$\mathbf{Tx} = \lambda_j \mathbf{x}, \quad \mathbf{Ty} = \lambda_k \mathbf{y}.$$

Consequently because $\lambda_j \neq \lambda_k$, we conclude that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal.

Let $\mathbb{M}$ be the subspace spanned by the union of the $\mathbb{H}_j$, $j \geq 0$. We claim that the closure of $\mathbb{M}$ is the full $\mathbb{H}$. Suppose, seeking a contradiction, that it is not so. Let $\mathbb{M}^\perp$ be the orthogonal complement to the subspace $\mathbb{M}$. Because $\mathbf{T}(\mathbb{M}) \subset \mathbb{M}$, if $\mathbf{y} \in \mathbb{M}^\perp$ and $\mathbf{x} \in \mathbb{M}$, then

$$\langle \mathbf{Ty}, \mathbf{x} \rangle = \langle \mathbf{y}, \mathbf{Tx} \rangle = 0.$$

The arbitrariness of $\mathbf{x}$ and $\mathbf{y}$ implies that $\mathbf{T}(\mathbb{M}^\perp) \subset \mathbb{M}^\perp$, as well. Thus $\mathbb{M}^\perp$ is a Hilbert space on its own, and $\mathbf{T}|_{\mathbb{M}^\perp}$ is a compact, self-adjoint operator. But since we have already recorded all of eigenvalues of $\mathbf{T}$ in the initial sequence $\{\lambda_j\}$, including $\lambda_0 = 0$, the restriction of $\mathbf{T}$ to $\mathbb{M}^\perp$ cannot have eigenvectors. According to the last item of Theorem 6.3, the only possibility left is for $\mathbb{M}^\perp$ to be trivial. Conclude then that $\overline{\mathbb{M}} = \mathbb{H}$ by Corollary 3.3.

If we organize over each $\mathbb{H}_j$ orthonormal bases, the union of them all will be an orthonormal basis for the full space $\mathbb{H}$, according to our proof here. The existence of orthonormal bases for $\mathbb{H}_0$ are guaranteed by Proposition 2.12 in case it is an infinite-dimensional subspace, while the $\mathbb{H}_j$'s, for $j \geq 1$, are finite dimensional. □

As in the finite-dimensional setting, this fundamental result is informing us that there is a special basis for $\mathbb{H}$, in which the compact, self-adjoint operator $\mathbf{T}$ admits a especially convenient representation: if

$$\mathbf{x} = \sum_j x_j \mathbf{e}_j, \quad \mathbf{Te}_j = \lambda_j \mathbf{e}_j,$$

i.e. $\{\mathbf{e}_j\}$ is an orthonormal basis of $\mathbb{H}$ with eigenvectors, then

$$\mathbf{Tx} = \sum_j \lambda_j x_j \mathbf{e}_j.$$

Example 6.6 Trigonometric bases. This is related to Example 2.14. Take $\mathbb{H}$ to be the Hilbert space $L^2(-\pi, \pi)$, and define the operator

$$\mathbf{T} : \mathbb{H} \to \mathbb{H}, \quad \mathbf{T}u(x) = \int_{-\pi}^{x} (y - x)u(y)\, dy - \frac{x + \pi}{2\pi} \int_{-\pi}^{\pi} (y - \pi)u(y)\, dy.$$

It is very easy to check that $\mathbf{T}$ is well-defined.

Moreover, it is compact because $\mathbf{T}$ is essentially an integral operator (check the examples in Sect. 6.1). It is very easy to realize that, in fact, if we put $U(x) = \mathbf{T}u(x)$ then

$$-U''(x) = u(x), \quad U(-\pi) = U(\pi) = 0,$$

i.e. U is the unique minimizer of the variational problem

$$\text{Minimize in } U(x) \in H_0^1(-\pi, \pi) : \quad \int_{-\pi}^{\pi} \left[\frac{1}{2} U'(x)^2 - u(x)U(x) \right] dx.$$

However, the image of $\mathbf{T}$ is contained in the subspace $L_0^2(-\pi, \pi)$ with vanishing average

$$\int_{-\pi}^{\pi} f(x)\, dx = 0.$$

This is very easily checked because this integral condition is correct for function $U \in H_0^1(-\pi, \pi)$. Restricted to this subspace $\mathbb{H} = L_0^2(\pi, \pi)$, $\mathbf{T}$ is also compact. $\mathbf{T}$ is self-adjoint too because if for $v(x) \in \mathbb{H}$, we put $V(x) = \mathbf{T}v(x)$, then

$$\langle \mathbf{T}u, v \rangle = \int_{-\pi}^{\pi} U(x)v(x)\, dx = -\int_{-\pi}^{\pi} U(x)V''(x)\, dx$$

and two successive integrations by parts yield

$$\langle \mathbf{T}u, v \rangle = -\int_{-\pi}^{\pi} U''(x)V(x)\, dx = \langle u, \mathbf{T}v \rangle.$$

According to Theorem 6.4, the full collection of, suitably normalized, eigenfunctions for $\mathbf{T}$ furnish an orthonormal basis for $L^2(-\pi, \pi)$. Such eigenfunctions, and their respective eigenvalues, are the posible non-trivial solutions of the problem

$$-U''(x) = \lambda U(x), \quad U(-\pi) = U(\pi) = 0.$$

It is an elementary exercise in Differential Equations to find that the only non-trivial solutions of this family of problems correspond to

$$\lambda = j^2/4, \quad j = 1, 2, \dots,$$

and

$$U_j(x) = \frac{1}{\sqrt{\pi}} \sin\left(\frac{1}{2} jx + \frac{1}{2} j\pi \right).$$

This family of trigonometric functions is an orthonormal basis for $L^2(-\pi, \pi)$. Note how different it is from the one in Example 2.14, in spite of both being orthonormal basis of the same space.

Example 6.7 A typical Hilbert-Schmidt operator. In the context of Example 6.5,

$$K(\mathbf{x}, \mathbf{y}) : \Omega \times \Omega \to \mathbb{R}, \quad \Omega \subset \mathbb{R}^N,$$

with

$$K(\mathbf{x}, \mathbf{y}) \in L^2(\Omega \times \Omega), \quad \int_{\Omega\times\Omega} |K(\mathbf{x}, \mathbf{y})|^2 \, d\mathbf{x} \, d\mathbf{y} < \infty,$$

define the corresponding integral operator

$$\mathbf{T}u(x) = \int_\Omega K(\mathbf{x}, \mathbf{y})u(\mathbf{y}) \, d\mathbf{y}, \quad \mathbf{T} : L^2(\Omega) \to L^2(\Omega).$$

If the kernel K is symmetric, then the operator $\mathbf{T}$ is compact and self-adjoint. As a consequence of Theorem 6.4, there is a basis of $L^2(\Omega)$ made up of eigenfunctions of $\mathbf{T}$. Eigenfunctions are solutions of the integral equation

$$\int_\Omega K(\mathbf{x}, \mathbf{y})u(\mathbf{y}) \, d\mathbf{y} = \lambda u(\mathbf{x}).$$

Of particular importance is the case of a convolution kernel

$$K(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{x} - \mathbf{y}).$$

It is hard to solve these integral equation even for simple situations.

Example 6.8 Definition of a function of an operator. If $\mathbf{T}$ is a compact, self-adjoint operator of a Hilbert space into itself, we already know there is a special orthonormal basis $\{\mathbf{e}_j\}$ very well adapted to $\mathbf{T}$ in such a way that if

$$\mathbf{x} = \sum_j x_j \mathbf{e}_j, \quad \mathbf{T}\mathbf{e}_j = \lambda_j \mathbf{e}_j, \tag{6.7}$$

then

$$\mathbf{T}\mathbf{x} = \sum_j \lambda_j x_j \mathbf{e}_j.$$

Given a real continuous function

$$f(\lambda) : \mathbb{R} \to \mathbb{R}, \quad f(0) = 0,$$

we can define the operator $f(\mathbf{T})$ which is a linear, compact, self-adjoint operator uniquely determined by putting

$$f(\mathbf{T})\mathbf{x} = \sum_j f(\lambda_j)x_j\mathbf{e}_j$$

under (6.7). See an explicit example below.

6.5 Exercises

1. Let $\mathbb{E}$, $\mathbb{F}$, $\mathbb{G}$, $\mathbb{H}$ be four Banach spaces. Suppose that

$$\mathbf{S} \in \mathcal{L}(\mathbb{E}, \mathbb{F}), \quad \mathbf{T} \in \mathcal{K}(\mathbb{F}, \mathbb{G}), \quad \mathbf{U} \in \mathcal{L}(\mathbb{G}, \mathbb{H}).$$

   Argue that the composition $\mathbf{U} \circ \mathbf{T} \circ \mathbf{S} \in \mathcal{K}(\mathbb{E}, \mathbb{H})$.
2. Show that if $\mathbb{U}$ and $\mathbb{M}$ are a bounded set and a subspace, respectively, in a Banach space $\mathbb{E}$, such that

$$\mathbb{U} \subset \mathbb{M} + \frac{1}{2}\mathbb{U},$$

   then $\mathbb{U} \subset \overline{\mathbb{M}}$.
3. Let $\mathbf{T} \in \mathcal{L}(\mathbb{H})$ with $\mathbb{H}$, a Hilbert space. Show that $\sigma(\mathbf{T}) = \sigma(\mathbf{T}')$, and that

$$\mathbb{N}(\mathbf{T}) = \mathbb{R}(\mathbf{T}')^{\perp}, \quad \mathbb{N}(\mathbf{T}') = \mathbb{R}(\mathbf{T})^{\perp}.$$

4. (a) Define the logarithm of a symmetric, positive-definite matrix $\mathbf{A}$. Calculate $\log \mathbf{A}$ for

$$\mathbf{A} = \begin{pmatrix} 3 & 1 & 1 \\ -4 & -1 & -2 \\ 2 & 1 & 2 \end{pmatrix}.$$

   (b) Argue that for the operator

$$\mathbf{T} : L^2(0, 1) \to L^2(0, 1), \quad U = \mathbf{T}u,$$
$$-U'' = u \text{ in } (0, 1), U(0) = U(1) = 0,$$

   its square root is well-defined.

5. Argue that there is a basis of $L^2(0, 1)$ composed of solutions of the problems

$$-u_j(x)'' + u_j(x) = \lambda_j u_j(x) \text{ in } (0, 1), \quad u_j'(0) = u_j'(1) = 0,$$

for a collection of numbers λ_j tending to ∞.

6. Let $\mathbb{E}$ be an infinite-dimensional, separable Hilbert space with

$$\mathbb{E} = \overline{\cup_i \mathbb{E}_i}, \quad \mathbb{E}_i \subset \mathbb{E}_{i+1},$$

and each $\mathbb{E}_i$ of finite dimension. Let π_i be the orthogonal projection onto $\mathbb{E}_i$, and $\pi_i^\perp$, the orthogonal projection onto its orthogonal complement. Argue that

$$\pi_i^\perp \mathbf{x} \to \mathbf{0}$$

for every $\mathbf{x}$, and yet it is not true that $\|\pi_i^\perp\| \to 0$. Does this fact contradict the Banach-Steinhaus principle?

7. Let $\mathbb{E}$ be a separable Hilbert space with

$$\mathbb{E} = \overline{\cup_i \mathbb{E}_i}, \quad \mathbb{E}_i \subset \mathbb{E}_{i+1},$$

and each $\mathbb{E}_i$ of finite dimension. Let π_i be the orthogonal projection onto $\mathbb{E}_i$. An operator $\mathbf{T} \in \mathcal{L}(\mathbb{E})$ is compact if and only if

$$\|\pi_i \mathbf{T} - \mathbf{T}\| \to 0, \quad i \to \infty.$$

8. Suppose $\mathbf{T} \in \mathcal{L}(\mathbb{H})$, $\mathbb{H}$ a Hilbert space, can be written

$$\mathbf{T}(\mathbf{x}) = \sum_{i=1}^{\infty} a_i \pi_i(\mathbf{x})$$

where π_i are the orthogonal projections onto finite-dimensional, pairwise-disjoint subspaces $\mathbf{K}_i$, and a_i are distinct, non-vanishing numbers converging to 0. Show that $\mathbf{T}$ is compact, commutes with its adjoint, the a_i's are its non-null eigenvalues, and

$$\mathbf{K}_i = \ker(\mathbf{T} - a_i \mathbf{1}).$$

9. Let $\mathbf{T} : \ell^2 \to \ell^2$ be defined by

$$\mathbf{T}(x_1, x_2, x_3, \dots) = (2x_1, x_3, x_2, 0, 0, \dots).$$

Prove that $\mathbf{T}$ is compact, and self-adjoint. Calculate $\sigma(\mathbf{T})$, the corresponding eigen-spaces, the orthogonal projections, and the spectral decomposition of $\mathbf{T}$.

10. For the kernel

$$K(x, y) = \begin{cases} (1-x)y, & 0 \le y \le x \le 1, \\ (1-y)x, & 0 \le x \le y \le 1, \end{cases}$$

consider the integral operator $\mathbf{T} : L^2(0,1) \to L^2(0,1)$ associated with K.

(a) If $\lambda \neq 0$ is an eigenvalue and u, a corresponding eigen-function, argue that u is C^∞ and

$$\lambda u'' + u = 0 \text{ in } (0,1), \quad u(0) = u(1) = 0.$$

(b) Conclude that

$$\sigma(\mathbf{T}) = \{(j\pi)^{-2} : j \in \mathbb{N}\} \cup \{0\}.$$

Find the associated eigen-functions. Is 0 an eigenvalue?

(c) Calculate $\mathbf{T}'$, and give conditions on v so that the equation

$$(\mathbf{T} - \lambda \mathbf{1})u = v$$

be solvable for $\lambda = (j\pi)^{-2}$.

11. Investigate whether the operators of Exercises 7 and 8 of Chap. 5 are compact.
12. For a continuous function $f(x) : [0,1] \to [0,1]$, define the operator

$$\mathbf{T}_f(u) = u(f(x)), \quad u \in C([0,1]).$$

Prove that $\mathbf{T}_f$ is compact if and only if f is constant.

13. Show that the integral equation

$$u(x) - \int_0^\pi \sin(x+y)u(y)\,dy = v(x), \quad x \in [0,\pi],$$

always has a solution for every $v \in C([0,\pi])$.

14. For a orthonormal basis $\{\mathbf{e}_j\}$ in a Hilbert space $\mathbb{H}$, let $\mathbf{T} \in \mathcal{L}(\mathbb{H})$ be such that

$$\sum_j \|\mathbf{T}\mathbf{e}_j\|^2 < \infty.$$

Prove that $\mathbf{T}$ is compact. Give an example of a compact operator $\mathbf{T}$ in ℓ^2 of the form

$$\mathbf{T}(\mathbf{x}) = (t_j x_j)_j, \quad \mathbf{x} = (x_1, x_2, \dots),$$

such that

$$\sum_j \|\mathbf{T}\mathbf{e}_j\|^2 = \infty, \quad \mathbf{e}_j = \chi_{\{j\}}.$$

15. Let $\mathbf{T}$ be a compact operator in a reflexive Banach space $\mathbb{E}$. Argue that the image of the unit ball $\mathbf{B}$ of $\mathbb{E}$ under $\mathbf{T}$ is closed. For the particular case $\mathbb{E} = C([-1, 1])$ and the Volterra operator

$$\mathbf{T}(u)(x) = \int_{-1}^{x} u(y)\,dy,$$

show that $\mathbf{T}(\mathbf{B})$ is not closed looking at the sequence

$$u_j(x) = \begin{cases} 0, & -1 \le x \le 0, \\ jx, & 0 \le x \le 1/j, \\ 1, & 1/j \le x \le 1 \end{cases}.$$

16. Argue that the injection operator

$$\iota : W^{1,p}(J) \to L^p(J), \quad \iota(u) = u$$

is a compact operator.
17. Use the Riesz lemma Lemma 6.1 to argue that if $\mathbb{E}$ is a normed space such that the unit ball $\mathbb{B}(\mathbf{0}, 1)$ is compact, then $\mathbb{E}$ is finite-dimensional.
18. Let $\mathbb{H} = L^2(0, 1)$, and

$$\mathbf{T} : \mathbb{H} \to \mathbb{H}, \quad \mathbf{T}f(s) = sf(s).$$

(a) Check that $\mathbf{T}$ is linear, continuous and self-adjoint with $\|\mathbf{T}\| = 1$.
(b) Show that $\sigma(\mathbf{T}) = [0, 1]$, and calculate $e(\mathbf{T})$. Is $\mathbf{T}$ compact?

19. The operator $\mathbf{T}$ determined by

$$f(x) \mapsto \mathbf{T}f(x) = \int_0^x f(y)\,dy - x\int_0^1 f(y)\,dy$$

yields the unique solution $u(x)$ of the problem

$$(u'(x) - f(x))' = 0 \text{ in } (0, 1), \quad u(0) = u(1) = 0.$$

(a) Show that the it is so.
(b) Prove that $\mathbf{T} : L^2(0, 1) \to L^2(0, 1)$ is compact.
(c) Show that the operator derivative

$$f(x) \mapsto \mathbf{T}' f(x) \equiv (\mathbf{T} f(x))'$$

is not compact.

Part III
Multidimensional Sobolev Spaces and Scalar Variational Problems

Chapter 7
Multidimensional Sobolev Spaces

7.1 Overview

We have introduced one-dimensional Sobolev spaces in Chap. 2. The fundamental concept that is required is that of weak derivative. In the case of higher dimension, we need to talk about weak partial derivatives. The basic identity that allows to introduce such fundamental concept is again the integration-by-parts formula which in the multidimensional framework is a consequence of the classical divergence theorem and the product rule. Though many concepts are similar to the one-dimensional situation, technical issues are much more involved.

Again, our initial motivation comes from our interest to deal with variational problems where one tries to minimize an integral functional of the form

$$I(u) = \int_{\Omega} F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x},$$

where $\Omega \subset \mathbb{R}^N$ is a certain domain; feasible functions

$$u(\boldsymbol{x}) : \Omega \to \mathbb{R}$$

should, typically, comply with further conditions like having preassigned boundary values around $\partial\Omega$; and the integrand

$$F(\boldsymbol{x}, u, \mathbf{u}) : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$$

determines in a crucial way the nature of the associated functional I.

Our first main task focuses on defining a weak gradient

$$\nabla u(\boldsymbol{x}) = \left(\frac{\partial u}{\partial x_1}(\boldsymbol{x}), \frac{\partial u}{\partial x_2}(\boldsymbol{x}), \dots, \frac{\partial u}{\partial x_N}(\boldsymbol{x})\right)$$

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*, La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4_7

that takes the place of the variable $\mathbf{u} \in \mathbb{R}^N$ in the integrand $F(x, u, \mathbf{u})$ when computing the cost $I(u)$ of every competing function u. Sobolev spaces will then be the class of all those functions whose weak derivates, i.e. weak gradient, enjoy certain integrability conditions. This chapter concentrates on setting the foundations for such Banach spaces, while the next chapter builds on Sobolev spaces to explore scalar, multidimensional, variational problems.

One of the main reasons why multidimensional Sobolev spaces are much more sophisticated than their one-dimensional counterpart, is that the class of domains in $\mathbb{R}^N$ where Sobolev spaces can be considered is infinitely more varied than in $\mathbb{R}$. On the other hand, there is no genuine multidimensional version of a fundamental theorem of Calculus as such, but every such result is based one way or another in the classical fundamental theorem of Calculus which is essentially one-dimensional.

The study of Sobolev spaces in the multidimensional setting is a quite reach and fascinating subject. Though in this chapter we focus our attention on those fundamental facts necessary for their use in minimization problems for standard variational problems, we will come back to them in the final chapter to explore some other crucial properties of functions belonging to these spaces to enlarge their use in variational problems and other important situations.

7.2 Weak Derivatives and Sobolev Spaces

Suppose

$$u(\boldsymbol{x}), \phi(\boldsymbol{x}) : \Omega \to \mathbb{R}$$

are two smooth functions, with ϕ having compact support within Ω. By the classical divergence theorem, we have

$$\int_\Omega \operatorname{div}[\phi(\boldsymbol{x})u(\boldsymbol{x})\mathbf{e}_i]\,d\boldsymbol{x} = \int_{\partial\Omega} \phi(\boldsymbol{x})u(\boldsymbol{x})\mathbf{e}_i \cdot \mathbf{n}(\boldsymbol{x})\,dS(\boldsymbol{x}), \tag{7.1}$$

where $\mathbf{n}(\boldsymbol{x})$ stands for the outer, unit normal to $\partial\Omega$, and $\{\mathbf{e}_i\}_{i=1,2,\dots,N}$ is the canonical basis of $\mathbb{R}^N$. The right-hand side of (7.1) vanishes because ϕ, having compact support contained in Ω, vanishes over $\partial\Omega$, while the left-hand side can be written, through the product rule, as

$$\int_\Omega [u(\boldsymbol{x})\nabla\phi(\boldsymbol{x}) \cdot \mathbf{e}_i + \phi(\boldsymbol{x})\nabla u(\boldsymbol{x}) \cdot \mathbf{e}_i]\,d\boldsymbol{x} = 0,$$

i.e.

$$\int_\Omega [u(\boldsymbol{x})\frac{\partial\phi}{\partial x_i}(\boldsymbol{x}) + \phi(\boldsymbol{x})\frac{\partial u}{\partial x_i}(\boldsymbol{x})]\,d\boldsymbol{x} = 0,$$

for every $i = 1, 2, \dots, N$. If we regard ϕ as a varying test function, this identity, written in the usual form

$$\int_\Omega u(\boldsymbol{x})\frac{\partial \phi}{\partial x_i}(\boldsymbol{x})\, d\boldsymbol{x} = -\int_\Omega \phi(\boldsymbol{x})\frac{\partial u}{\partial x_i}(\boldsymbol{x})\, d\boldsymbol{x}$$

permits us to declare a function $u_i(\boldsymbol{x})$ as the weak ith-partial derivative of u if

$$\int_\Omega u(\boldsymbol{x})\frac{\partial \phi}{\partial x_i}(\boldsymbol{x})\, d\boldsymbol{x} = -\int_\Omega \phi(\boldsymbol{x})u_i(\boldsymbol{x})\, d\boldsymbol{x}$$

for every smooth, compactly supported ϕ. The coherence of such a concept is clear, for if u is indeed smooth with a gradient (in the usual sense) $\nabla u(\boldsymbol{x})$ then

$$\nabla u(\boldsymbol{x}) = \mathbf{u}(\boldsymbol{x}), \quad \mathbf{u} = (u_1, u_2, \dots, u_N).$$

But there are non-differentiable functions u in the classical sense which admit weak gradients. We write $\nabla u(\boldsymbol{x})$ to designate the vector of weak, partial derivatives $\mathbf{u}(\boldsymbol{x})$ of u, and then we recover the formula of integration by parts

$$\int_\Omega u(\boldsymbol{x})\nabla\phi(\boldsymbol{x})\, d\boldsymbol{x} = -\int_\Omega \phi(\boldsymbol{x})\nabla u(\boldsymbol{x})\, d\boldsymbol{x} \tag{7.2}$$

which is valid for every smooth function ϕ with compact support in Ω.

Mimicking the one-dimensional situation, we introduce the multidimensional Sobolev spaces.

Definition 7.1 Let Ω be an open subset of $\mathbb{R}^N$, and let exponent $p \in [1, +\infty]$ be given. The Sobolev space $W^{1,p}(\Omega)$ is defined as the collection of functions of $L^p(\Omega)$ admitting a weak gradient ∇u, in the above sense, which is a vector with components in $L^p(\Omega)$ as well. In compact form, we write

$$W^{1,p}(\Omega) = \{u \in L^p(\Omega) : \nabla u \in L^p(\Omega; \mathbb{R}^N)\}.$$

The norm in $W^{1,p}(\Omega)$ is

$$\|u\|_{W^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \|\nabla u\|_{L^p(\Omega;\mathbb{R}^N)}.$$

Equivalently, we can take, in a more explicit way,

$$\|u\|^p_{W^{1,p}(\Omega)} = \int_\Omega [|u(\boldsymbol{x})|^p + |\nabla u(\boldsymbol{x})|^p]\, d\boldsymbol{x}.$$

As in the one-dimensional situation, the exponent $p = 2$ is a very special case. To highlight it, we use a special notation, which is universally accepted, and put

$$H^1(\Omega) \equiv W^{1,2}(\Omega).$$

There is an inner product in $H^1(\Omega)$ given by

$$\langle u, v\rangle = \int_\Omega [u(\boldsymbol{x})v(\boldsymbol{x}) + \nabla u(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x})]\, d\boldsymbol{x},$$

with associated norm

$$\|u\|^2_{H^1(\Omega)} = \int_\Omega [u^2(\boldsymbol{x}) + |\nabla u(\boldsymbol{x})|^2]\, d\boldsymbol{x}.$$

Our first statement asserts that these spaces are Banach spaces. In addition, $H^1(\Omega)$ is a Hilbert space.

Proposition 7.1 $W^{1,p}(\Omega)$ *is a Banach space for every exponent* $p \in [1, +\infty]$. $H^1(\Omega)$ *is a separable Hilbert space.*

The proof of this fact is exactly like the one for Proposition 2.3.

Remark 7.1 Most of the basic properties of multi-dimensional Sobolev spaces are inherited from the corresponding Lebesgue spaces. Features like reflexivity and separability hold for $W^{1,p}(\Omega)$ if $p \in (1, +\infty)$, while $W^{1,1}(\Omega)$ is also separable. When one works with weak derivatives, duality is not that important. In fact, there is no standard way to identify the dual of $W^{1,p}(\Omega)$. It is more important to identify the dual spaces of certain subspaces of $W^{1,p}(\Omega)$ as we will see. On the other hand, we have not paid much attention to the separability property except for Hilbert spaces.

Since even just measurable functions are nearly continuous according to Luzin's theorem, when deciding in practice if a given function belongs to a certain Sobolev space, one would differentiate such function paying attention to places where such differentiation is not legitimate, and proceed from there. However, it may be not so easy to finally decide if a given function belongs to a Sobolev space in the multi-dimensional scenario because the set of singularities of derivatives may be quite intricate and its size, and integrability properties, at various dimensions may be crucial in the final decision. This may be a quite delicate and specialized issue well beyond the scope of this text. There are however easier situations in which one can decide in a more direct way.

Proposition 7.2 *Let* $u(\boldsymbol{x})$ *be a function in* $L^\infty(\Omega)$ *that is continuous and differentiable except in a singular subset* ω *of* Ω *of measure zero that admits a family of regular subsets* Ω_ϵ *(in the sense that the divergence theorem is valid in* $\Omega \setminus \Omega_\epsilon$*) with*

$$\omega \subset \Omega_\epsilon, \quad |\partial\Omega_\epsilon|, |\Omega_\epsilon| \to 0$$

as $\epsilon \searrow 0$*, and its gradient* $\nabla u(\boldsymbol{x}) \in L^p(\Omega)$ *for some* $p \in [1, \infty]$*. Then* $u \in W^{1,p}(\Omega)$*, and its weak gradient is* $\nabla u(\boldsymbol{x})$*.*

Proof Let $\phi(\boldsymbol{x})$ be a smooth, compactly-supported function in Ω, and take the family of subsets Ω_ϵ indicated in the statement such that u is continuous and differentiable in $\Omega \setminus \Omega_\epsilon$ and $|\partial\Omega_\epsilon|, |\Omega_\epsilon| \searrow 0$. By our hypotheses, the three integrals

$$\int_{\Omega_\epsilon} \phi(\boldsymbol{x})\nabla u(\boldsymbol{x})\,d\boldsymbol{x}, \quad \int_{\partial\Omega_\epsilon} \phi(\boldsymbol{x})u(\boldsymbol{x})\mathbf{n}(\boldsymbol{x})\,dS(\boldsymbol{x}),$$
$$\int_{\Omega_\epsilon} \nabla\phi(\boldsymbol{x})u(\boldsymbol{x})\,d\boldsymbol{x},$$

converge to zero as $\epsilon \searrow 0$. The first one because it can be bounded from above by

$$\left(\int_{\Omega_\epsilon} |\phi(\boldsymbol{x})|^q\,d\boldsymbol{x}\right)^{1/q} \left(\int_{\Omega_\epsilon} |\nabla u(\boldsymbol{x})|^p\,d\boldsymbol{x}\right)^{1/p},$$

and this integral converges to zero if $|\nabla u|^p$ is integrable and $|\Omega_\epsilon| \searrow 0$ (this is even so if $p = 1$). The second and third ones are straightforward if $u \in L^\infty(\Omega)$ is continuous.

On the other hand, by the regularity properties assumed on u and on $\Omega \setminus \Omega_\epsilon$, we can apply the divergence theorem in $\Omega \setminus \Omega_\epsilon$, and find

$$\int_{\Omega\setminus\Omega_\epsilon} \nabla\phi(\boldsymbol{x})u(\boldsymbol{x})\,d\boldsymbol{x} = -\int_{\Omega\setminus\Omega_\epsilon} \phi(\boldsymbol{x})\nabla u(\boldsymbol{x})\,d\boldsymbol{x} + \int_{\partial\Omega_\epsilon} \phi(\boldsymbol{x})u(\boldsymbol{x})\mathbf{n}(\boldsymbol{x})\,dS(\boldsymbol{x}),$$

where $\mathbf{n}(\boldsymbol{x})$ is the unit, "inner" normal to the hypersurface $\partial\Omega_\epsilon$.

With these facts, let us examine the following chain of identities

$$\begin{aligned}
\int_{\Omega} \nabla\phi(\boldsymbol{x})u(\boldsymbol{x})\,d\boldsymbol{x} &= \lim_{\epsilon\to 0}\left(\int_{\Omega\setminus\Omega_\epsilon} \nabla\phi(\boldsymbol{x})u(\boldsymbol{x})\,d\boldsymbol{x} + \int_{\Omega_\epsilon} \nabla\phi(\boldsymbol{x})u(\boldsymbol{x})\,d\boldsymbol{x}\right)\\
&= \lim_{\epsilon\to 0}\int_{\Omega\setminus\Omega_\epsilon} \nabla\phi(\boldsymbol{x})u(\boldsymbol{x})\,d\boldsymbol{x}\\
&= \lim_{\epsilon\to 0}\left(-\int_{\Omega\setminus\Omega_\epsilon} \phi(\boldsymbol{x})\nabla u(\boldsymbol{x})\,d\boldsymbol{x}\right.\\
&\qquad \left.+\int_{\partial\Omega_\epsilon} \phi(\boldsymbol{x})u(\boldsymbol{x})\mathbf{n}(\boldsymbol{x})\,dS(\boldsymbol{x})\right)\\
&= -\lim_{\epsilon\to 0}\int_{\Omega\setminus\Omega_\epsilon} \phi(\boldsymbol{x})\nabla u(\boldsymbol{x})\,d\boldsymbol{x}
\end{aligned}$$

$$= - \lim_{\epsilon \to 0} \left(\int_{\Omega \setminus \Omega_\epsilon} \phi(\boldsymbol{x}) \nabla u(\boldsymbol{x}) \, d\boldsymbol{x} + \int_{\Omega_\epsilon} \phi(\boldsymbol{x}) \nabla u(\boldsymbol{x}) \, d\boldsymbol{x} \right)$$

$$= - \int_{\Omega} \phi(\boldsymbol{x}) \nabla u(\boldsymbol{x}) \, d\boldsymbol{x}.$$

The arbitrariness of ϕ in this final identity implies our claim. □

Example 7.1 The family of functions

$$u_i(\boldsymbol{x}) = \frac{x_i}{|\boldsymbol{x}|}, \quad i = 1, 2, \ldots, N,$$

are the components of the mapping

$$\mathbf{u}(\boldsymbol{x}) = \frac{\boldsymbol{x}}{|\boldsymbol{x}|}.$$

They are differentiable except at the origin. Off this point, the functions are smooth, and their gradients are easily calculated as

$$\nabla \mathbf{u}(\boldsymbol{x}) = \frac{1}{|\boldsymbol{x}|} \mathbf{1} - \frac{1}{|\boldsymbol{x}|^3} \boldsymbol{x} \otimes \boldsymbol{x}, \quad (\boldsymbol{x} \otimes \boldsymbol{x})_{ij} = x_i x_j, \boldsymbol{x} = (x_i)_i,$$

where, as usual, $\mathbf{1}$ is the identity matrix of the appropriate dimension ($N \times N$). It is then elementary to realize that

$$|\nabla \mathbf{u}(\boldsymbol{x})| \leq \frac{C}{|\boldsymbol{x}|},$$

for some positive constant C. We therefore see, according to Proposition 7.2 where Ω_ϵ is the ball $\mathbf{B}_\epsilon$ of radius ϵ around the origen, that the Sobolev space to which these functions belong to depend on the integrability of the function

$$u(\boldsymbol{x}) = \frac{1}{|\boldsymbol{x}|^p}.$$

Lemma 7.1 *For $R > 0$, the integral*

$$\int_{\mathbf{B}_R} |z|^s \, dz, \quad \mathbf{B}_R = \{\boldsymbol{x} \in \mathbb{R}^N : |\boldsymbol{x}| \leq 1\},$$

is finite whenever $s > -N$.

According to this elementary lemma, we conclude that our family of functions in this example belong to $W^{1,p}(\Omega)$ for a bounded domain containing the origen at least for $1 \leq p < N$; and if the origen does not belong to $\overline{\Omega}$, for every p.

Example 7.2 For a bounded domain $\Omega \subset \mathbb{R}^2$ with non-empty intersection with the Y-axis, consider the function

$$u(x_1, x_2) = \frac{x_1}{|x_1|}.$$

It is clear that the second partial derivative vanishes; however, the first one has a jump all along the part J of the Y-axis within Ω. The point is that this set J cannot be covered by a sequence like the one in Proposition 7.2. Indeed, it is impossible to have that $\partial\Omega_\epsilon$ be made arbitrarily small; after all, the length, the 1-dimensional measure, of J does not vanish. This function u does not belong to any Sobolev space, but to some more general space whose study goes beyond the goal of this text.

7.3 Completion of Spaces of Smooth Functions of Several Variables with Respect to Integral Norms

As in Sect. 2.10, we can talk about the vector space $C^\infty(\Omega)$ of smooth functions in a open subset $\Omega \subset \mathbb{R}^N$. Endowed with the integral norm

$$\|u\|^p = \int_\Omega \left(|u(\boldsymbol{x})|^p + |\nabla u(\boldsymbol{x})|^p\right]), d\boldsymbol{x}, \quad u \in C^\infty(\Omega),\ p \geq 1,$$

the space is not complete. But it can be immediately completed according to Theorem 2.1. The space so obtained is then complete with respect to the above norm. In it, smooth functions are dense.

We could legitimately call the resulting space $W^{1,p}(\Omega)$, and argue that functions u in this space admit a weak gradient $\nabla u \in L^p(\Omega; \mathbb{R}^N)$. What we cannot know in advance is whether this space is the full set of functions in $L^p(\Omega)$ with such a weak gradient. This would require to work harder to be convinced that smooth functions are dense, under this norm, in the set of all functions in $L^p(\Omega)$ with a weak gradient in $L^p(\Omega; \mathbb{R}^N)$, much in the same way as in Sect. 2.29. Under additional assumptions on Ω, this density fact can be shown to hold. See Corollary 9.1 below.

From the viewpoint of our interest in variational problems, one can work in the completion of $C^\infty(\Omega)$ with respect to the above norm, and proceed from there. Nothing essentially changes in the analysis that we will perform in the next chapter.

7.4 Some Important Examples

We describe in this section some important examples of multi-dimensional Sobolev functions of special relevance.

Suppose we have a measurable kernel

$$K(\boldsymbol{x}, \boldsymbol{y}) : \Omega \times \Omega \to \mathbb{R}$$

with the fundamental property that

$$K(\cdot, \boldsymbol{y}) \in W^{1,1}(\Omega)$$

for a.e. $\boldsymbol{y} \in \Omega$. In addition, there are symmetric, measurable, non-negative kernels

$$K_0(\boldsymbol{x}, \boldsymbol{y}), \quad K_1(\boldsymbol{x}, \boldsymbol{y})$$

with

$$|K(\boldsymbol{x}, \boldsymbol{y})| \le K_0(\boldsymbol{x}, \boldsymbol{y}), \quad |\nabla_{\boldsymbol{x}} K(\boldsymbol{x}, \boldsymbol{y})| \le K_1(\boldsymbol{x}, \boldsymbol{y}),$$

and the measurable functions

$$\mu_i(\boldsymbol{x}) = \int_\Omega K_i(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y}, \quad i = 0, 1, \tag{7.3}$$

turn out to be bounded in Ω (they belong to $L^\infty(\Omega)$). For a given measurable function $f(\boldsymbol{y})$, define

$$u(\boldsymbol{x}) = \int_\Omega K(\boldsymbol{x}, \boldsymbol{y}) f(\boldsymbol{y})\, d\boldsymbol{y}, \quad \boldsymbol{x} \in \Omega, \tag{7.4}$$

through the corresponding Hilbert-Schmidt type operator.

Lemma 7.2 *If $f(\boldsymbol{y}) \in L^p(\Omega)$ for some $p \ge 1$, then $u(\boldsymbol{x})$ defined through (7.4), under all of the previous assumptions, belongs to $W^{1,p}(\Omega)$, and*

$$\nabla u(\boldsymbol{x}) = \int_\Omega \nabla_{\boldsymbol{x}} K(\boldsymbol{x}, \boldsymbol{y}) f(\boldsymbol{y})\, d\boldsymbol{y}. \tag{7.5}$$

Proof We first check that $u(\boldsymbol{x})$ defined through (7.4) belongs to $L^p(\Omega)$. We first have

$$|u(\boldsymbol{x})| \le \int_\Omega |K(\boldsymbol{x}, \boldsymbol{y})|\, |f(\boldsymbol{y})|\, d\boldsymbol{y} \le \int_\Omega K_0(\boldsymbol{x}, \boldsymbol{y})\, |f(\boldsymbol{y})|\, d\boldsymbol{y}.$$

Since $\mu_0(\boldsymbol{x})$ is finite for a.e. $\boldsymbol{x} \in \Omega$, according to (7.3), the measure

$$d\nu_{\boldsymbol{x}}(\boldsymbol{y}) = \frac{K_0(\boldsymbol{x}, \boldsymbol{y})}{\mu_0(\boldsymbol{x})}\, d\boldsymbol{y}$$

is a probability measure. By Jensen's inequality applied to the function $|\cdot|^p$, which is convex if $p \geq 1$,

$$\left| \int_\Omega |f(\boldsymbol{y})| \, d\nu_{\boldsymbol{x}}(\boldsymbol{y}) \right|^p \leq \int_\Omega |f(\boldsymbol{y})|^p \, d\nu_{\boldsymbol{x}}(\boldsymbol{y})$$
$$= \frac{1}{\mu_0(\boldsymbol{x})} \int_\Omega K_0(\boldsymbol{x}, \boldsymbol{y}) \, |f(\boldsymbol{y})|^p \, d\boldsymbol{y}.$$

Hence

$$|u(\boldsymbol{x})|^p \leq \mu_0(\boldsymbol{x})^{p-1} \int_\Omega K_0(\boldsymbol{x}, \boldsymbol{y}) \, |f(\boldsymbol{y})|^p \, d\mathbf{y}.$$

If we integrate in $\boldsymbol{x} \in \Omega$, we arrive at

$$\|u\|^p_{L^p(\Omega)} \leq \|\mu_0\|^{p-1}_{L^\infty(\Omega)} \int_\Omega \int_\Omega K_0(\boldsymbol{x}, \boldsymbol{y}) \, |f(\boldsymbol{y})|^p \, d\boldsymbol{y} \, d\boldsymbol{x}.$$

By Fubini's theorem, and the symmetry of K_0, we can also write

$$\|u\|^p_{L^p(\Omega)} \leq \|\mu_0\|^p_{L^\infty(\Omega)} \, \|f\|^p_{L^p(\Omega)}.$$

Concerning derivatives, it is immediate to check formula (7.5) testing it against a smooth function $\phi(\boldsymbol{x})$ with compact support in Ω as in (7.2). In checking this, it is also crucial to use Fubini's theorem to interchange the order of integration. Finally each weak partial derivative in (7.5) belongs to $L^p(\Omega)$, by the same ideas as with (7.4) replacing the kernel $K_0(\boldsymbol{x}, \boldsymbol{y})$ by $K_1(\boldsymbol{x}, \boldsymbol{y})$. □

One fundamental particular case follows. Define the newtonian potential $w(\boldsymbol{x})$ of an integrable function $f(\boldsymbol{y}) : \Omega \to \mathbb{R}$ by the formula

$$w(\boldsymbol{x}) = \int_\Omega \Theta(\boldsymbol{x} - \boldsymbol{y}) f(\boldsymbol{y}) \, d\boldsymbol{y}, \quad \boldsymbol{x} \in \mathbb{R}^N, \tag{7.6}$$

where

$$\Theta(\boldsymbol{x} - \boldsymbol{y}) = \Theta(|\boldsymbol{x} - \boldsymbol{y}|) = \begin{cases} \frac{1}{N(2-N)\omega_N} |\boldsymbol{x} - \boldsymbol{y}|^{2-N}, & N > 2, \\ \frac{1}{2\pi} \log |\boldsymbol{x} - \boldsymbol{y}|, & N = 2. \end{cases}$$

The positive number ω_N is the measure of the unit ball in $\mathbb{R}^N$, and $N\omega_N$ the surface measure of the unit sphere. Note that w in (7.6) is defined for every $\boldsymbol{x} \in \mathbb{R}^N$, though

$f(\boldsymbol{y})$ is only defined for $\boldsymbol{y} \in \Omega$. It is true however that we can also rewrite (7.6) in the form

$$w(\boldsymbol{x}) = \int_{\mathbb{R}^N} \Theta(\boldsymbol{x} - \boldsymbol{y})\chi_\Omega(\boldsymbol{y}) f(\boldsymbol{y})\, d\boldsymbol{y}$$

for the characteristic function $\chi_\Omega(\boldsymbol{y})$ defined as 1 for points $\boldsymbol{y} \in \Omega$, and vanishing off Ω.

According to our general discussion above, we need to check that

$$\int_\Omega \Theta(\boldsymbol{x} - \boldsymbol{y})\, d\boldsymbol{y}$$

is uniformly bounded in $\boldsymbol{x} \in \Omega$, to conclude that $w(\boldsymbol{x})$ defined through (7.6) belongs to $L^p(\Omega)$ if f belongs to this same space. Notice that we can take

$$K_0(\boldsymbol{x}, \boldsymbol{y}) = \Theta(\boldsymbol{x} - \boldsymbol{y}).$$

Now, if Ω is bounded, and $\Omega - \Omega \subset \mathbf{B}_R$ for some positive $R > 0$, then

$$\begin{aligned} \int_\Omega \Theta(\boldsymbol{x} - \boldsymbol{y})\, d\boldsymbol{y} =& \frac{1}{N(2-N)\omega_N} \int_{\boldsymbol{x}-\Omega} |z|^{2-N}\, dz \\ &\le \frac{1}{N(2-N)\omega_N} \int_{\mathbf{B}_R} |z|^{2-N}\, dz, \end{aligned}$$

and this last integral is finite according to Lemma 7.1, and independent of $\boldsymbol{x} \in \Omega$. The case $N = 2$ can also be checked separately. Consequently, if $f \in L^p(\Omega)$, so does $w(\boldsymbol{x})$ in (7.6). But more is true.

Lemma 7.3 *If $f \in L^p(\Omega)$ with Ω, bounded, and $p \ge 1$, then its newtonian potential $w(\boldsymbol{x})$ given by (7.6) belongs to $W^{1,p}(\Omega)$.*

Proof After our previous calculations, all we need to find is a symmetric, upper bound $K_1(\boldsymbol{x}, \boldsymbol{y})$

$$|\nabla_{\boldsymbol{x}} \Theta(\boldsymbol{x} - \boldsymbol{y})| \le K_1(\boldsymbol{x}, \boldsymbol{y})$$

with

$$\int_\Omega K_1(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y}$$

bounded uniformly in $\boldsymbol{x} \in \Omega$. It is an elementary exercise to find that

$$\nabla_{\boldsymbol{x}} \Theta(\boldsymbol{x} - \boldsymbol{y}) = \frac{1}{N\omega_N} |\boldsymbol{x} - \boldsymbol{y}|^{-N}(\boldsymbol{x} - \boldsymbol{y}),$$

and hence, we can take

$$K_1(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{N\omega_N}|\boldsymbol{x} - \boldsymbol{y}|^{1-N}.$$

Lemma 7.1 enables us to conclude that

$$\int_\Omega K_1(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y}$$

is uniformly bounded in $\boldsymbol{x} \in \Omega$, and hence $w(\boldsymbol{x}) \in W^{1,p}(\Omega)$. □

Remark 7.2 One can explore the possibility of a further differentiation. Indeed, we are led to compute the hessian matrix of $w(\boldsymbol{x})$ in (7.6) through the formula

$$\nabla^2_{\boldsymbol{x}}\Theta(\boldsymbol{x} - \boldsymbol{y}) = \frac{1}{N\omega_N|\boldsymbol{x} - \boldsymbol{y}|^N}\left(\mathbf{1}_N - N\frac{\boldsymbol{x} - \boldsymbol{y}}{|\boldsymbol{x} - \boldsymbol{y}|} \otimes \frac{\boldsymbol{x} - \boldsymbol{y}}{|\boldsymbol{x} - \boldsymbol{y}|}\right),$$

where $\mathbf{1}_N$ is the identity matrix of size $N \times N$, and, as already introduced earlier,

$$\mathbf{u} \otimes \mathbf{v} = (u_i v_j)_{ij}, \quad \mathbf{u} = (u_i) \in \mathbb{R}^N, \mathbf{v} = (v_j) \in \mathbb{R}^N$$

is the tensor product of two vectors in $\mathbb{R}^N$. Hence

$$|\nabla^2_{\boldsymbol{x}}\Theta(\boldsymbol{x} - \boldsymbol{y})| \leq K_2(\boldsymbol{x}, \boldsymbol{y}) \equiv C_N|\boldsymbol{x} - \boldsymbol{y}|^{-N}$$

for a certain constant C_N. However, the integral

$$\int_\Omega K_2(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y}$$

is no longer finite according to Lemma 7.1. We can easily check, though, that

$$\Delta_{\boldsymbol{x}}\Theta(\boldsymbol{x} - \boldsymbol{y}) = \mathrm{tr}(\nabla^2_{\boldsymbol{x}}\Theta(\boldsymbol{x} - \boldsymbol{y})) = 0$$

for every $\boldsymbol{y}$ except when $\boldsymbol{x} = \boldsymbol{y}$. We will come back to these facts in the final chapter.

Another interesting example is the generalization of the results in Sect. 2.29 to a higher dimensional situation. Specifically, take a smooth function $\rho(\boldsymbol{x}) : \mathbb{R}^N \to \mathbb{R}$, supported in the unit ball $\mathbf{B}$, and with

$$\int_{\mathbf{B}} \rho(\boldsymbol{x})\, d\boldsymbol{x} = 1.$$

The typical choice is similar to the one in the above-mentioned section

$$\rho(\boldsymbol{x}) = \begin{cases} C \exp \frac{1}{|\boldsymbol{x}|^2-1}, & |\boldsymbol{x}| \leq 1, \\ 0, & |\boldsymbol{x}| \geq 1, \end{cases}$$

for a suitable positive constant C. Define

$$K(\boldsymbol{x}, \boldsymbol{y}) = \rho(\boldsymbol{x} - \boldsymbol{y}), \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N,$$

and

$$u(\boldsymbol{x}) = \int_{\mathbb{R}^N} \rho(\boldsymbol{x} - \boldsymbol{y}) f(\boldsymbol{y})\, d\boldsymbol{y} \tag{7.7}$$

for $f \in L^p(\mathbb{R}^N)$. As a direct corollary of Lemma 7.2, based on the smoothness of ρ, the following is immediate.

Corollary 7.1

1. *For $f \in L^p(\mathbb{R}^N)$, the function u in (7.7) belongs to the space $W^{1,p}(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$.*
2. *Moreover if*

$$\rho_j(\boldsymbol{z}) = \frac{1}{j}\rho(j\boldsymbol{z}), \quad f_j(\boldsymbol{x}) = \int_{\mathbb{R}^N} \rho_j(\boldsymbol{x} - \boldsymbol{y}) f(\boldsymbol{y})\, d\boldsymbol{y}$$

for $f \in W^{1,p}(\mathbb{R}^N)$, then $f_j \to f$ in $W^{1,p}(\mathbb{R}^N)$.

The proof is similar to the one for Lemma 2.4.

7.5 Domains for Sobolev Spaces

Sobolev spaces can be defined for any open set $\Omega \subset \mathbb{R}^N$. Yet, from the viewpoint of variational problems, fundamental properties of Sobolev functions, that one cannot be dispensed with, can only be shown under some reasonable properties of the sets were Sobolev functions are defined and considered. There are three main facts that Sobolev functions should comply with to be of use in variational problems of interest with sufficient generality:

1. Sobolev functions in $W^{1,p}(\Omega)$ should have traces on the boundary $\partial\Omega$, i.e. typical variational problems demand preassigned boundary values for competing functions;
2. coercivity often asks, under boundary conditions, for the $L^p(\Omega)$-norm of a function in $W^{1,p}(\Omega)$ to be included in the $L^p(\Omega)$-norm of its gradient (Poincaré's inequality);

3. weak convergence of derivatives must imply strong convergence of functions: the injection from $W^{1,p}(\Omega)$ into $L^p(\Omega)$ must be a compact operator.

There are various ways to formulate geometric properties on subsets Ω for which the three above properties hold. We have adopted the one we believe is most economical in terms of technicalities. It makes multidimensional Sobolev spaces to raise naturally, and in a transparent form, from their one-dimensional version, as we focus on the restriction of functions to one-dimensional fibers or sections.

One initial technical lemma will be helpful. Curiously, it is usually known as the Fundamental Lemma of the Calculus of Variations.

Lemma 7.4 *Let $\Phi(\boldsymbol{x})$ be a measurable, locally integrable function defined in an open set $D \subset \mathbb{R}^N$. If for every $\phi(\boldsymbol{x})$ smooth, and with compact support in D, we have that*

$$\int_D \Phi(\boldsymbol{x})\phi(\boldsymbol{x})\, d\boldsymbol{x} = 0, \tag{7.8}$$

then $\Phi(\boldsymbol{x}) = 0$ for a.e. $\boldsymbol{x} \in D$.

Proof The proof is straightforward if we are allowed to use general measure-theoretic results. Other possibilities use approximation by mollifiers to generalize (7.8) for functions $\phi \in L^\infty(D)$ as in Lemma 2.4. We seek a contradiction between (7.8), and the fact that Φ might not vanish somewhere. Suppose we could find some non-negligible set $\mathbf{C}$ where Φ is strictly positive. We can approximate $\mathbf{C}$ by non-negligible compact subsets $\mathbf{K}$ from within. If then ϕ is selected strictly positive and with compact support in $\mathbf{K}$, we would find that the integral in (7.8) would be strictly positive, a contradiction. The same argument works if Φ is negative in a non-negligible subset. □

The following proposition is the clue to our definition below of a domain. It clearly explains why Sobolev functions in $W^{1,p}(\Omega)$ are in fact Sobolev functions in one-dimensional fibers in a generalized cylinder Ω.

Proposition 7.3 *Let Γ be an open piece of the hyperplane $x_N = 0$ in $\mathbb{R}^N$, and $J_{\boldsymbol{x}'}$, an open interval in $\mathbb{R}$ for a.e. $\boldsymbol{x}' \in \Gamma$ such that*

$$\Omega = \Gamma + J_{\boldsymbol{x}'}\mathbf{e}_N = \{\boldsymbol{x}' + t\mathbf{e}_N : \boldsymbol{x}' \in \Gamma, t \in J_{\boldsymbol{x}'}\} \tag{7.9}$$

is an open subset (a cylinder with $\mathbf{e}_N$-axis and variable vertical fibers) of $\mathbb{R}^N$. If $u \in W^{1,p}(\Omega)$, the function of one variable

$$t \in J_{\boldsymbol{x}'} \mapsto u(\boldsymbol{x}' + t\mathbf{e}_N) = u(\boldsymbol{x}', t)$$

belongs to $W^{1,p}(J_{\boldsymbol{x}'})$ for a.e. $\boldsymbol{x}' \in \Gamma$, and hence it is absolutely continuous (as a function of one variable).

Proof If $u(\boldsymbol{x}) \in W^{1,p}(\Omega)$, then

$$\int_{\Omega} \left[u(\boldsymbol{x}) \frac{\partial \phi}{\partial x_N}(\boldsymbol{x}) + \phi(\boldsymbol{x}) \frac{\partial u}{\partial x_N}(\boldsymbol{x}) \right] d\boldsymbol{x} = 0 \tag{7.10}$$

for every test function $\phi(\boldsymbol{x})$, smooth and compactly-supported in Ω. In particular, we can take test functions of the product form

$$\phi(\boldsymbol{x}) = \phi'(\boldsymbol{x}')\phi_N(\boldsymbol{x}', x_N), \quad \boldsymbol{x} = (\boldsymbol{x}', x_N), \boldsymbol{x}' \in \Gamma,$$

for arbitrary ϕ', smooth and compactly-supported in Γ. By taking smooth test functions of this form into (7.10), we realize that

$$\int_{\Gamma} \phi'(\boldsymbol{x}') \int_{J_{\boldsymbol{x}'}} \left[u(\boldsymbol{x}) \frac{\partial \phi_N}{\partial x_N}(\boldsymbol{x}) + \phi_N(\boldsymbol{x}) \frac{\partial u}{\partial x_N}(\boldsymbol{x}) \right] dx_N \, d\boldsymbol{x}' = 0.$$

If we put

$$\Phi(\boldsymbol{x}') \equiv \int_{J_{\boldsymbol{x}'}} \left[u(\boldsymbol{x}', x_N) \frac{\partial \phi_N}{\partial x_N}(\boldsymbol{x}', x_N) + \phi_N(\boldsymbol{x}', x_N) \frac{\partial u}{\partial x_N}(\boldsymbol{x}', x_N) \right] dx_N,$$

by the preceding lemma, we conclude that

$$\int_{J_{\boldsymbol{x}'}} \left[u(\boldsymbol{x}', x_N) \frac{\partial \phi_N}{\partial x_N}(\boldsymbol{x}', x_N) + \phi_N(\boldsymbol{x}', x_N) \frac{\partial u}{\partial x_N}(\boldsymbol{x}', x_N) \right] dx_N = 0$$

vanishes for a.e. $\boldsymbol{x}' \in \Gamma$. The arbitrariness of ϕ_N implies our result. Note that for fixed $\boldsymbol{x}' \in \Gamma$, the test function $\phi_N(\boldsymbol{x}', x_N)$ can be taken to be of the product form too (Exercise 1). □

All we need for this proof to be valid is that the last partial derivative $\partial u/\partial x_N$ belong to $L^p(\Omega)$.

We expect our definition below will be clearly sensed after the preceding result. We designate by $\{\mathbf{e}_i\}$, the canonical basis of $\mathbb{R}^N$ and by π_i, the i-th coordinate projection of $\mathbb{R}^N$ onto $\mathbb{R}^{N-1}$, $i = 1, 2, \dots, N$.

Definition 7.2 We will say that an open subset $\Omega \subset \mathbb{R}^N$ is a domain if it enjoys the following "cylinder" property:

> There is a finite number n, independent of i, such that for every $i = 1, 2, \dots, N$, and for every $\boldsymbol{x}' \in \pi_i\Omega$, there is $J_{i,\boldsymbol{x}'} \subset \mathbb{R}$ which is a finite union of at most n open intervals (some of which could share end-points), with
>
> $$\Omega = \pi_i\Omega + J_{i,\boldsymbol{x}'}\mathbf{e}_i = \{\boldsymbol{x}' + t\mathbf{e}_i : \boldsymbol{x}' \in \pi_i\Omega, t \in J_{i,\boldsymbol{x}'}\}, \tag{7.11}$$
>
> for every $i = 1, 2, \dots, N$.

Any reasonable set will fall under the action of this definition. Singular sets violating this condition are related to not having the finiteness of n.

Proposition 7.3 can be strengthened in the sense that partial derivatives can be measured along any orthogonal system of coordinates. Once functions posses weak partial derivatives with respect to one such system, then they have weak partial derivatives with respect to every other such system too.

Proposition 7.4 *Let $\Omega \subset \mathbb{R}^N$ be a domain, and $u \in W^{1,p}(\Omega)$. For a.e. pair of points $\boldsymbol{x}, \boldsymbol{y} \in \Omega$, the one-dimensional section*

$$f(t) : J_{\boldsymbol{x},\boldsymbol{y}} \subset \mathbb{R} \to \mathbb{R}, \quad J_{\boldsymbol{x},\boldsymbol{y}} = \{t \in \mathbb{R} : t\boldsymbol{x} + (1-t)\boldsymbol{y} \in \Omega\},$$
$$f(t) = u(t\boldsymbol{x} + (1-t)\boldsymbol{y}),$$

is absolutely continuous, and for a.e. $t \in [0, 1]$,

$$f'(t) = \nabla u(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}).$$

Proof For an arbitrary rotation $\mathbf{R}$ in $\mathbb{R}^N$, $\mathbf{R}^T\mathbf{R} = \mathbf{1}$, define the function

$$u_{\mathbf{R}}(z) : \mathbf{R}^T\Omega \to \mathbb{R}, \quad u_{\mathbf{R}}(z) = u(\mathbf{R}z).$$

It is easy to check that this new function $u_{\mathbf{R}}$ belongs to $W^{1,p}(\mathbf{R}^T\Omega)$ (Exercise 3), and that

$$\nabla u_{\mathbf{R}}(z) = \mathbf{R}^T\nabla u(\mathbf{R}z).$$

Proposition 7.3 can then be applied to $u_{\mathbf{R}}$ to conclude, due to the arbitrariness of $\mathbf{R}$, that u is also absolutely continuous along a.e. (one-dimensional) line intersecting Ω. $\square$

In practice, quite often domains are usually defined through functions in such a way that the regularity or smoothness of such functions determined the regularity of their associated domains.

Definition 7.3 We will say that a domain $\Omega \subset \mathbb{R}^N$ is a C^1-domain if there is a C^1-function $\phi(\boldsymbol{x}) : \mathbb{R}^N \to \mathbb{R}$ such that

1. $\Omega = \{\phi > 0\}, \quad \mathbb{R}^N \setminus \Omega = \{\phi < 0\}, \quad \partial\Omega = \{\phi = 0\}$;
2. for some positive ϵ, the function ϕ has no critical point in the strip $\{-\epsilon < \phi < \epsilon\}$ around $\partial\Omega$.

The domain

$$\Omega_\epsilon = \{-\epsilon < \phi\}$$

is an extension of Ω. The normal direction to $\partial\Omega$ is given by the gradient $\nabla\phi$, which, by hypothesis, is non-singular over $\partial\Omega$. Every regular domain according to Definition 7.3 is a domain according to Definition 7.2.

If Ω is a C^1-domain, the signed-distance function to $\partial\Omega$ is one standard choice for ϕ

$$\phi(\boldsymbol{x}) = \operatorname{dist}(\boldsymbol{x}, \partial\Omega).$$

A ball is, of course, a prototypical case

$$\mathbf{B}_R(\boldsymbol{x}_0) = \{\boldsymbol{x} \in \mathbb{R}^N : |\boldsymbol{x} - \boldsymbol{x}_0|^2 < R\}, \quad \phi(\boldsymbol{x}) = R - |\boldsymbol{x} - \boldsymbol{x}_0|^2.$$

The definition of a domain through a smooth function ϕ permits some interesting manipulations with functions $u \in W^{1,p}(\Omega)$ as we will later see. For instance, the following fact may be interesting in some circumstances.

Lemma 7.5 *Let $\Omega \subset \mathbb{R}^N$ be a C^1-domain with bounded boundary determined by $\phi(\boldsymbol{x})$. Then there is a sequence $\{\eta_j\}$ of C^1-functions defined in Ω such that:*

1. $\eta_j(\mathbf{x}) : \Omega \to [0, 1]$, $\eta_j = 0$ *on* $\partial\Omega$*;*
2. *if* $\omega_j = |\Omega \setminus \{\eta_j = 1\}|$, *then* $\omega_j \searrow 0$ *as* $j \to \infty$*;*
3. $|\nabla\eta_j| \le C/\omega_j$, *for some constant C (that may depend on Ω).*

Proof Select a C^1-function $P(y)$ in $[0, 1]$ with the four properties

$$P(0) = P'(0) = P'(1) = 0, \quad P(1) = 1,$$

for instance the polynomial

$$P(y) = 3y^2 - 2y^3.$$

Extend it by putting

$$P(y) = 1 \text{ for } y \ge 1, \quad P(y) = 0 \text{ for } y \le 0.$$

The resulting function is C^1 in all of $\mathbb{R}$. For a positive integer j, put

$$P_j(y) = P(jy), \quad \eta_j(x) = P_j(\phi(\boldsymbol{x})).$$

It is clear that each P_j is C^1, and that

$$\{P_j = 0\} = \mathbb{R}^N \setminus (\Omega \cup \partial\Omega), \quad \{P_j = 1\} = \{\phi \ge 1/j\},$$

at least for large j. In this way, by the standard co-area formula or Cavalieri's principle,

$$
\begin{aligned}
\omega_j =&|\Omega \setminus \{\eta_j = 1\}| \\
=&|\{0 < \phi \leq 1/j\}| \\
=&\int_0^{1/j} |\{\phi = s\}|\, ds.
\end{aligned}
$$

The smoothness of ϕ and the boundedness of $\partial\Omega$ imply that indeed $\omega_j \searrow 0$. Finally,

$$
\nabla\eta_j(\boldsymbol{x}) = P_j'(\phi(\boldsymbol{x}))\nabla\phi(\boldsymbol{x}).
$$

This last product vanishes except when $0 < \phi < 1/j$. For points $\boldsymbol{x}$ in this region, we find

$$
\nabla\eta_j(\boldsymbol{x}) = jP'(j\phi(\boldsymbol{x}))\nabla\phi(\boldsymbol{x}).
$$

Every factor in this product is bounded by a constant, independent of j, except the factor j itself. But the relationship between j and ω_j involves a further constant because ω_j is of the order of $1/j$. □

There is no particular difficulty in raising the degree of smoothness in Ω and in ϕ, though one would have to replace the polynomial P by a better auxiliary function. Sometimes one may need a slight different version of these cut-off functions $\{\eta_j\}$, which require a slight modification of the above proof.

Lemma 7.6 *Let $\Omega \subset \mathbb{R}^N$ be a C^1-domain with bounded boundary determined by $\phi(\boldsymbol{x})$. Then there is a sequence $\{\eta_j\}$ of C^1-functions defined in Ω such that*

1. *there is compact set $\mathbf{K} \subset \Omega$, such that $\operatorname{supp}(\eta_j) \subset \mathbf{K}$ for all j;*
2. *if $\omega_j = |\Omega \setminus \{\eta_j = 1\}|$, then $\omega_j \searrow 0$ as $j \to \infty$;*
3. *$|\nabla\eta_j| \leq C/\omega_j$, for some constant C (that may depend on Ω).*

We next explore the validity of the three main features we need for Sobolev functions to be of use in variational problems.

7.6 Traces of Sobolev Functions: The Space $W_0^{1,p}(\Omega)$

One vital ingredient in variational problems is the boundary condition around $\partial\Omega$ imposed on competing functions. This forces us to examine in what sense functions in Sobolev spaces can take on boundary values. It is another important consequence of Proposition 7.3, and our definition of a feasible domain Definition 7.2. Note that

functions in $L^p(\Omega)$ cannot have, in general, traces over $N-1$ dimensional sets in $\mathbb{R}^N$, as these have a vanishing N-dimensional Lebesgue measure.

We need first an interesting technical result involving the boundary $\partial\Omega$ of a domain to facilitate proofs, which is implicit in the definition of a domain. In fact, this property can be included in Definition 7.2 as part of it.

Lemma 7.7 *Let Ω be a domain according to Definition 7.2. Then for a.e. point $\boldsymbol{x}$ in the boundary $\partial\Omega$, there is some (most likely more than one) $i \in \{1, 2, \dots, N\}$, possibly depending on $\boldsymbol{x}$, such that*

$$\boldsymbol{x} = \pi_i \boldsymbol{x} + t\mathbf{e}_i, \quad t \in \overline{J_{i,\pi_i \boldsymbol{x}}}. \tag{7.12}$$

Proof In the context of Definition 7.2, consider the set

$$\cup_i \left(\pi_i\Omega + \partial J_{i,\boldsymbol{x}'}\mathbf{e}_i\right) = \cup_i \{\boldsymbol{x}' + t\mathbf{e}_i : \boldsymbol{x}' \in \pi_i\Omega, t \in \partial J_{i,\boldsymbol{x}'}\}$$

which is a subset of $\partial\Omega$. Each of the sets, for fixed i,

$$\{\boldsymbol{x}' + t\mathbf{e}_i : \boldsymbol{x}' \in \pi_i\Omega, t \in \partial J_{i,\boldsymbol{x}'}\}$$

covers the part of $\partial\Omega$ that is projected onto $\pi_i\Omega$, and hence our result is proved once we realize that the set

$$\partial\Omega \setminus \cup_i \left(\pi_i^{-1}(\pi_i\Omega) \cap \partial\Omega\right)$$

is negligible in $\partial\Omega$. □

Our main result in this section follows.

Proposition 7.5 *Let $\Omega \subset \mathbb{R}^N$ be a domain, and $u \in W^{1,p}(\Omega)$. Then the restriction*

$$u|_{\partial\Omega} : \partial\Omega \to \mathbb{R}$$

is well-defined, and measurable.

Proof Through Lemma 7.7, consider a partition of $\partial\Omega$ in N subsets Γ_i, $i = 1, 2, \dots, N$, so that for every $\boldsymbol{x} \in \Gamma_i$, (7.12) holds. In this way, we can restrict attention to the case in which the decomposition (7.12) holds for a particular i, and show that Sobolev functions in $W^{1,p}(\Omega)$ have a trace for a.e. $\boldsymbol{x} \in \Gamma_i$. Consider the case, without loss of generality, $i = N$, and put

$$\Gamma = \Gamma_N, \quad \boldsymbol{x}' = \pi_N \boldsymbol{x}, \quad J_{\boldsymbol{x}'}, \text{ a connected subinterval of } J_{N,\pi_N \boldsymbol{x}}.$$

We can write through Proposition 7.3, for every pair of numbers y and z in $J_{\boldsymbol{x}'}$,

$$u(\boldsymbol{x}', y) - u(\boldsymbol{x}', z) = \int_z^y \frac{\partial u}{\partial x_N}(\boldsymbol{x}', x)\, dx$$
$$= \int_{J_{\boldsymbol{x}'}} \chi_{[z,y]}(x) \frac{\partial u}{\partial x_N}(\boldsymbol{x}', x)\, dx,$$

and, by Hölder's inequality if $p > 1$,

$$|u(\boldsymbol{x}', y) - u(\boldsymbol{x}', z)| \le \left(\int_{J_{\boldsymbol{x}'}} \left| \frac{\partial u}{\partial x_N}(\boldsymbol{x}', x) \right|^p dx \right)^{1/p} |y - z|^{1/q}, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Since the quantities

$$\left(\int_{J_{\boldsymbol{x}'}} \left| \frac{\partial u}{\partial x_N}(\boldsymbol{x}', x) \right|^p dx \right)^{1/p}$$

are finite for a.e. $\boldsymbol{x}' \in \Omega'$ (because the integral of its p-th power is finite), we conclude that u is absolutely continuous along a.e. such fiber, and hence it is well-defined on a.e. point of the form

$$\boldsymbol{x}' + t\mathbf{e}_N, \quad t \in \overline{J_{\boldsymbol{x}'}}.$$

This implies our conclusion over $\Gamma = \Gamma_N$ since points in this set are exactly of this form. The case $p = 1$ is argued in the same way, though the above Hölder's inequality is not valid. □

Once we have shown that Sobolev functions in $W^{1,p}(\Omega)$, for a certain domain Ω, have traces on the boundary $\partial\Omega$, one can isolate the following important subspace.

Definition 7.4 For a domain $\Omega \subset \mathbb{R}^N$, the subspace

$$W_0^{1,p}(\Omega) \subset W^{1,p}(\Omega)$$

is the closure in $W^{1,p}(\Omega)$ of the subspace $C_c^\infty(\Omega)$ of smooth functions with compact support in Ω.

It is evident that the set of smooth functions with compact support in Ω have a vanishing trace on $\partial\Omega$. In fact, one can, equivalently, define $W_0^{1,p}(\Omega)$ as the class of functions with a vanishing trace on $\partial\Omega$. If we adopt the above definition, however, this must be a result to be proved.

Proposition 7.6 *Every function in $W_0^{1,p}(\Omega)$ has a vanishing trace on $\partial\Omega$.*

Proof It consists in the realization that the proof above of Proposition 7.5 is valid line by line for a full sequence $\{u_j\}$ converging in $W^{1,p}(\Omega)$, in such a way that we conclude that there is a point-wise convergence of their traces around $\partial\Omega$. If these vanish for a sequence of compactly supported functions, so does the trace of the limit function. □

There still remains the issue of whether every function in $W^{1,p}(\Omega)$ with a vanishing trace on $\partial\Omega$ belongs, as a matter of fact, to $W_0^{1,p}(\Omega)$. This amounts to showing that such functions can be approximated, in the norm of $W^{1,p}(\Omega)$, by a sequence of smooth functions with compact support contained in Ω. This issue is technical and requires some smoothness on $\partial\Omega$, as in other situations. Though we treat below some other similar points, we will take for granted, under the appropriate smoothness on Ω, that $W_0^{1,p}(\Omega)$ is exactly the subspace of $W^{1,p}(\Omega)$ with a vanishing trace on $\partial\Omega$. Under this equivalence, we find that (7.2) is correct for $u \in W^{1,p}(\Omega)$ and $\phi \in W_0^{1,q}(\Omega)$. In particular, it is correct for $u, \phi \in H^1(\Omega)$ and one of the two in $H_0^1(\Omega)$.

The most important point is that $W_0^{1,p}(\Omega)$ is a Banach space on its own right, under the same norm. $H_0^1(\Omega)$ is a Hilbert space with the same inner product of $H^1(\Omega)$. It suffices to check that $W_0^{1,p}(\Omega)$ is a closed subspace in $W^{1,p}(\Omega)$.

Proposition 7.7 *$W_0^{1,p}(\Omega)$ is closed in $W^{1,p}(\Omega)$.*

Proof Suppose we are facing a situation where

$$u_j \to u \text{ in } W^{1,p}(\Omega), \quad u_j \in W_0^{1,p}(\Omega),$$

and we would like to conclude that necessarily $u \in W_0^{1,p}(\Omega)$.

By Definition 7.4, we know that there is a sequence $\{\phi_j\}$ of smooth functions with compact support contained in Ω and such that

$$\|\phi_j - u_j\| \to 0 \text{ as } j \to \infty \text{ in } W^{1,p}(\Omega).$$

On the other hand

$$\|u_j - u\| \to 0 \text{ as } j \to \infty \text{ in } W^{1,p}(\Omega).$$

It is immediate to conclude that

$$\|\phi_j - u\| \to 0 \text{ as } j \to \infty \text{ in } W^{1,p}(\Omega),$$

and $u \in W_0^{1,p}(\Omega)$, again by Definition 7.4. □

Another natural and relevant question is what functions, defined in the boundary $\partial\Omega$ of a domain $\Omega \subset \mathbb{R}^N$, can be attained as the trace of a Sobolev function in Ω. There is a very precise answer to this question that is a bit beyond this first course

about Sobolev spaces. A practical answer that is sufficient in practice most of the time, is that such traces functions are, of course, of the form

$$u|_{\partial\Omega}, \quad u \in W^{1,p}(\Omega).$$

In this way, fixed boundary values around $\partial\Omega$ for a certain variational problem are given by providing a specific Sobolev function $u_0 \in W^{1,p}(\Omega)$, and then feasible functions $u \in W^{1,p}(\Omega)$ are asked to comply with the requirement

$$u - u_0 \in W_0^{1,p}(\Omega).$$

Remark 7.3 Note that $W_0^{1,p}(\mathbb{R}^N) = W^{1,p}(\mathbb{R}^N)$.

The following is another remarkable but natural result.

Proposition 7.8

1. *If $u \in W_0^{1,p}(\Omega)$, then its extension by zero, indicated by the operator* $\bar{\cdot}$,

$$\overline{u}(\boldsymbol{x}) = \begin{cases} u(\boldsymbol{x}), & \boldsymbol{x} \in \Omega, \\ 0, & \boldsymbol{x} \notin \Omega, \end{cases}$$

is a function in $W^{1,p}(\mathbb{R}^N)$. Moreover

$$\nabla\overline{u}(\boldsymbol{x}) = \begin{cases} \nabla u(\boldsymbol{x}), & \boldsymbol{x} \in \Omega, \\ \mathbf{0}, & \boldsymbol{x} \notin \Omega, \end{cases}$$

and

$$\|u\|_{W_0^{1,p}(\Omega)} = \|\overline{u}\|_{W^{1,p}(\Omega)}.$$

2. *If Ω is a regular domain (according to Definition 7.3), and the extended function $\overline{u}$ belongs to $W^{1,p}(\mathbb{R}^N)$, then $u \in W_0^{1,p}(\Omega)$.*

Proof Suppose first that $u \in W_0^{1,p}(\Omega)$, and let $u_j \in C_c^\infty(\Omega)$ such that

$$\|u - u_j\|_{W^{1,p}(\Omega)} \to 0.$$

It is evident that $\overline{u}_j \in C_c^\infty(\mathbb{R}^N)$, and that

$$\|\overline{u} - \overline{u}_j\|_{W^{1,p}(\mathbb{R}^N)} = \|u - u_j\|_{W^{1,p}(\Omega)} \to 0.$$

Hence $\overline{u} \in W^{1,p}(\mathbb{R}^N)$. This part does not require any smoothness on Ω.

Conversely, assume that $\overline{u} \in W^{1,p}(\mathbb{R}^N)$. By Remark 7.3, there is a sequence $\{u_j\}$ of smooth functions with compact support such that

$$\|\overline{u} - u_j\|_{W^{1,p}(\mathbb{R}^N)} \to 0.$$

But since

$$\|u_j\|_{W^{1,p}(\mathbb{R}^N \setminus \Omega)} \leq \|\overline{u} - u_j\|_{W^{1,p}(\mathbb{R}^N)},$$

we conclude that $u_j \to 0$ in $W^{1,p}(\mathbb{R}^N \setminus \Omega)$. If Ω is regular and ϕ is its defining function, by Lemma 7.6 (or rather its C^∞-version) there is a sequence $\{\eta_j\}$ with those stated properties. Then the sequence $\{v_j = u_j \eta_j\} \subset C_c^\infty(\Omega)$ and

$$\|u_j - v_j\|_{W^{1,p}(\Omega)} \to 0.$$

Therefore

$$\begin{aligned}\|u - v_j\|_{W^{1,p}(\Omega)} \leq& \|u - u_j\|_{W^{1,p}(\Omega)} + \|u_j - v_j\|_{W^{1,p}(\Omega)} \\ \leq& \|\overline{u} - u_j\|_{W^{1,p}(\mathbb{R}^N)} + \|u_j - v_j\|_{W^{1,p}(\Omega)},\end{aligned}$$

and $u \in W_0^{1,p}(\Omega)$. □

7.7 Poincaré's Inequality

Our manipulations in the proof of Proposition 7.5 lead in a natural way to the following remarkable fact. Recall that π_i, $i = 1, 2, \ldots, N$, is the i-th canonical coordinate projection so that $\mathbf{1} - \pi_i$ is the projection onto the i-th axis.

Proposition 7.9 *Suppose $\Omega \subset \mathbb{R}^N$ is a domain such that at least one of the N projections $(\mathbf{1} - \pi_i)\Omega$ is a bounded set of $\mathbb{R}$. Then there is a constant $C > 0$ (depending on p and on the size of this projection in $\mathbb{R}$) such that, for every $u \in W_0^{1,p}(\Omega)$, we have*

$$\|u\|_{L^p(\Omega)} \leq C \|\nabla u\|_{L^p(\Omega;\mathbb{R}^N)}.$$

Proof Suppose, without loss of generality, that the index i is the last one N, so that the diameter of the projection $(\mathbf{1} - \pi_N)\Omega$ onto the last axis is not greater than $L > 0$. We can write, with the notation in the proof of Proposition 7.5 above, for a.e. $\boldsymbol{x}' \in \pi_N \Omega$,

$$u(\boldsymbol{x}', y) = \int_z^y \frac{\partial u}{\partial x_N}(\boldsymbol{x}', x)\, dx = \int_{J_{\boldsymbol{x}'}} \chi_{[z,y]}(x) \frac{\partial u}{\partial x_N}(\boldsymbol{x}', x)\, dx,$$

if the point $(\boldsymbol{x}', z) \in \partial\Omega$, and hence $u(\boldsymbol{x}', z) = 0$. The diameter of the set $J_{\boldsymbol{x}'}$ is not greater than L for a.e. $\boldsymbol{x}' \in \pi_N\Omega$. Again by Hölder's inequality,

$$|u(\boldsymbol{x}', y)| \le \left(\int_{J_{\boldsymbol{x}'}} \left|\frac{\partial u}{\partial x_N}(\boldsymbol{x}', x)\right|^p \, dx\right)^{1/p} |y - z|^{1/q}, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

that is to say

$$|u(\boldsymbol{x}', y)| \le \left(\int_{J_{\boldsymbol{x}'}} \left|\frac{\partial u}{\partial x_N}(\boldsymbol{x}', x)\right|^p \, dx\right)^{1/p} L^{1/q},$$

for all $y \in J_{\boldsymbol{x}'}$. Therefore

$$\int_{J_{\boldsymbol{x}'}} |u(\boldsymbol{x}', y)|^p \, dy \le \int_{J_{\boldsymbol{x}'}} \left|\frac{\partial u}{\partial x_N}(\boldsymbol{x}', x)\right|^p \, dx \, L^{p/q+1}.$$

A further integration with respect to $\boldsymbol{x}'$ leads to

$$\begin{aligned}
\|u\|^p_{L^p(\Omega)} &= \int_{\pi_N\Omega} \int_{J_{\boldsymbol{x}'}} |u(\boldsymbol{x}', y)|^p \, dy \, d\boldsymbol{x}' \\
&\le L^p \int_{\pi_N\Omega} \int_{J_{\boldsymbol{x}'}} \left|\frac{\partial u}{\partial x_N}(\boldsymbol{x}', x)\right|^p \, dx \, d\boldsymbol{x}' \\
&= L^p \left\|\frac{\partial u}{\partial x_N}\right\|^p_{L^p(\Omega)}.
\end{aligned}$$

It is then clear that

$$\|u\|_{L^p(\Omega)} \le L \|\nabla u\|_{L^p(\Omega;\mathbb{R}^N)}.$$

Notice that $\frac{p}{q} + 1 = p$. The case $p = 1$ is also correct, and the proof requires some very minor adjustments. □

This result indicates that when boundary values around Ω are preassigned, the size of functions is somehow incorporated in the norm of the gradient. In particular, we see that the p-th norm of the gradient

$$\|\nabla u\|^p_{L^p(\Omega)} = \int_\Omega |\nabla u(\boldsymbol{x})|^p \, d\boldsymbol{x}$$

is truly a norm in the space $W_0^{1,p}(\Omega)$, and

$$\int_\Omega \nabla u(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x})\, d\boldsymbol{x}$$

is a genuine inner product in $H_0^1(\Omega)$. Poincaré's inequality is the second main point that we set to ourselves before proceeding to examining scalar, multidimensional variational problems.

7.8 Weak and Strong Convergence

One main ingredient to reduce in an important way the convexity requirement on integrands for multidimensional variational problems, is to show that weak convergence in $W^{1,p}(\Omega)$ implies strong convergence in $L^p(\Omega)$. We will focus on these properties from a broader perspective in the final chapter, but now we will treat this issue in a right-to-the-point manner. Our plan is to complete Corollary 2.4.

Proposition 7.10 *Let $p > 1$, and suppose $\{u_j\}$ is a bounded sequence of functions in $W^{1,p}(\Omega)$ for a domain $\Omega \subset \mathbb{R}^N$. There is a subsequence, not relabeled, and a function $u \in W^{1,p}(\Omega)$ such that $u_j \rightharpoonup u$ in $W^{1,p}(\Omega)$, and $u_j \to u$ in $L^p(\Omega)$.*

Proof The N sequences of functions

$$\{\partial u_j/\partial x_i\}, \quad i = 1, 2, \dots, N,$$

are bounded in $L^p(\Omega)$. By the first part of Corollary 2.4, there are functions

$$u^{(i)} \in L^p(\Omega), \quad i = 0, 1, 2, \dots, N,$$

with

$$\frac{\partial u_j}{\partial x_i} \rightharpoonup u^{(i)}, i = 1, 2, \dots, N, \quad u_j \rightharpoonup u^{(0)}.$$

We first claim that $u \equiv u^{(0)}$ belongs to $W^{1,p}(\Omega)$, and its i-th weak, partial derivative is precisely $u^{(i)}$. To this end, take a test function ϕ, and write for each j and i, since $u_j \in W^{1,p}(\Omega)$,

$$\int_\Omega [u_j(\boldsymbol{x}) \frac{\partial \phi}{\partial x_i}(\boldsymbol{x}) + \phi(\boldsymbol{x}) \frac{\partial u_j}{\partial x_i}(\boldsymbol{x})]\, d\boldsymbol{x} = 0.$$

By the claimed weak convergences, a direct passage to the limit in j, leads to

$$\int_\Omega [u(\boldsymbol{x})\frac{\partial \phi}{\partial x_i}(\boldsymbol{x}) + \phi(\boldsymbol{x})u^{(i)}(\boldsymbol{x})]\,d\boldsymbol{x} = 0.$$

This identity, valid for arbitrary test functions ϕ, exactly means, given that each $u^{(i)}$ belongs to $L^p(\Omega)$, that $u \in W^{1,p}(\Omega)$, $u^{(i)}$ is the i-th partial derivative of u, and hence, $u_j \rightharpoonup u$ in $W^{1,p}(\Omega)$.

It remains to show the fundamental additional fact that $u_j \to u$ (strong) in $L^p(\Omega)$. We know that

$$\int_{\Omega'}\int_{J_{\boldsymbol{x}'}} \left|\frac{\partial u_j}{\partial x_N}(\boldsymbol{x}', x_N)\right|^p dx_N\,d\boldsymbol{x}' = \left\|\frac{\partial u_j}{\partial x_N}\right\|^p \le M < \infty,$$

$$\frac{\partial u_j}{\partial x_N} \rightharpoonup \frac{\partial u}{\partial x_N},$$

for a positive constant M, independent of j. This weak convergence means that

$$\int_\Omega \left(\frac{\partial u_j}{\partial x_N}(\boldsymbol{x}) - \frac{\partial u}{\partial x_N}(\boldsymbol{x})\right)\phi(\boldsymbol{x})\,d\boldsymbol{x} \to 0 \tag{7.13}$$

for all $\phi \in L^q(\Omega)$. By (7.11) in Definition 7.2

$$\Omega = \Omega' \times \{J_{\boldsymbol{x}'}\mathbf{e}_N : \boldsymbol{x}' \in \Omega'\}, \qquad \Omega' = \pi_N\Omega \subset \mathbb{R}^{N-1},$$

and we can recast (7.13) in the form

$$\int_{\Omega'}\int_{J_{\boldsymbol{x}'}} \left(\frac{\partial u_j}{\partial x_N}(\boldsymbol{x}', x_N) - \frac{\partial u}{\partial x_N}(\boldsymbol{x}', x_N)\right)\phi(\boldsymbol{x}', x_N)\,dx_N\,d\boldsymbol{x}' \to 0.$$

In particular, we can take ϕ of the product form

$$\phi(\boldsymbol{x}', x_N) = \psi(\boldsymbol{x}')\phi(x_N),$$

to find

$$\int_{\Omega'} \psi(\boldsymbol{x}')\left(\int_{J_{\boldsymbol{x}'}} \left(\frac{\partial u_j}{\partial x_N}(\boldsymbol{x}', x_N) - \frac{\partial u}{\partial x_N}(\boldsymbol{x}', x_N)\right)\phi(x_N)\,dx_N\right) d\boldsymbol{x}' \to 0.$$

Due to the arbitrariness of ψ, thanks to Lemma 7.4, we can conclude that, for a.e. $\boldsymbol{x}' \in \Omega'$,

$$\frac{\partial u_j}{\partial x_N}(\boldsymbol{x}', \cdot) \in L^p(J_{\boldsymbol{x}'}), \qquad \frac{\partial u_j}{\partial x_N}(\boldsymbol{x}', \cdot) \rightharpoonup \frac{\partial u}{\partial x_N}(\boldsymbol{x}', \cdot).$$

From Proposition 2.7 and the observations before its statement, we can conclude that

$$u_j(\boldsymbol{x}', \cdot) + v_j(\boldsymbol{x}') \to u(\boldsymbol{x}', \cdot)$$

for certain measurable functions v_j, independent of x_N, and for a.e. $\boldsymbol{x}' \in \Omega'$. For this fact to be precisely true, one would have to partition the domain Ω in subsets where the transversal sets $J_{\boldsymbol{x}'}$ of $\mathbb{R}$ are single intervals (Exercise 2 below). Our conclusion exactly means that we have the point-wise convergence

$$u_j(\boldsymbol{x}) + v_j(\boldsymbol{x}') \to u(\boldsymbol{x})$$

for a.e. $\boldsymbol{x} \in \Omega$, and a.e. $\boldsymbol{x}' \in \Omega'$. There is nothing keeping us from going over this argument with a different partial derivative

$$\partial/\partial x_i, \quad i = 1, 2, \dots, N-1,$$

so that we would conclude that in fact the functions $v_j(\boldsymbol{x}')$ can be taken as constants v_j independent of $\boldsymbol{x}$. Since, on the other hand, we indeed know that $u_j \rightharpoonup u$ in $L^p(\Omega)$, by uniqueness of limits, we conclude that $v_j \to 0$, because weak and strong convergence for constants is the same, and $u_j \to u$ strongly in $L^p(\Omega)$. □

With this result we complete our initial analysis of first-order, multidimensional Sobolev spaces that permits us to deal with the most pertinent issues about scalar variational problems. We will do so in the next chapter. We include a final section to briefly describe how to set up, in an inductive manner, higher-order Sobolev spaces. These will allow to deal with higher-order variational problems.

7.9 Higher-Order Sobolev Spaces

Once we have defined Sobolev spaces of first-order involving first-order weak derivatives belonging to Lebesgue spaces, it is easy to move forward and define second-order Sobolev spaces, and high-order Sobolev spaces.

Definition 7.5 Let Ω be an open subset of $\mathbb{R}^N$, and let exponent $p \in [1, +\infty]$ be given. The Sobolev space $W^{2,p}(\Omega)$ is defined as the collection of functions in $W^{1,p}(\Omega)$ whose partial derivatives

$$\frac{\partial u}{\partial x_i} \in L^p(\Omega), \quad i = 1, 2, \dots, N,$$

admit a weak gradient

$$\nabla \frac{\partial u}{\partial x_i} = \left(\frac{\partial}{\partial x_1}\frac{\partial u}{\partial x_i}, \frac{\partial}{\partial x_2}\frac{\partial u}{\partial x_i}, \dots, \frac{\partial}{\partial x_N}\frac{\partial u}{\partial x_i} \right).$$

As in the smooth case, the full collection of weak second partial derivatives can be arranged in the weak hessian

$$\nabla^2 u = \left(\frac{\partial^2 u}{\partial x_i \partial x_j} \right)_{i,j=1,2,\dots,N}$$

which is always a symmetric $N \times N$-matrix. In compact form, we write

$$W^{2,p}(\Omega) = \{u \in L^p(\Omega) : \nabla u \in L^p(\Omega; \mathbb{R}^N), \nabla^2 u \in L^p(\Omega; \mathbb{R}^{N\times N})\}.$$

$W^{2,p}(\Omega)$ is a Banach space under the norm

$$\|u\|^p = \int_\Omega \left(|u(\boldsymbol{x})|^p + |\nabla u(\boldsymbol{x})|^p + |\nabla^2 u(\boldsymbol{x})|^p \right) d\boldsymbol{x},$$

or, equivalently,

$$\|u\| = \|u\|_{L^p(\Omega)} + \|\nabla u\|_{L^p(\Omega;\mathbb{R}^N)} + \|\nabla^2 u\|_{L^p(\Omega:\mathbb{R}^{N\times N})}.$$

The space $H^2(\Omega) = W^{2,2}(\Omega)$ is a separable, Hilbert space with the inner product

$$\langle u, v\rangle = \int_\Omega \left(u(\boldsymbol{x})v(\boldsymbol{x}) + \nabla u(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) + \nabla^2 u(\boldsymbol{x}) : \nabla^2 v(\boldsymbol{x}) \right) d\boldsymbol{x}.$$

In a similar, inductive way, one can define Sobolev spaces $W^{m,p}(\Omega)$ for $m \geq 1$, by demanding that derivatives of order $m-1$ belong to $W^{1,p}(\Omega)$. Recall that the product $\mathbf{A} : \mathbf{B}$ of two $N \times N$-matrices is, as usual,

$$\mathbf{A} : \mathbf{B} = \mathrm{tr}(\mathbf{A}^T \mathbf{B}).$$

The fact that the weak hessian is a symmetric matrix, because of the equality of mixed partial derivatives, is also a direct consequence of the same fact for smooth functions through the integration-by-parts formula.

Most of the important facts for functions in $W^{2,p}(\Omega)$ can be deduced from parallel facts for functions in $W^{1,p}(\Omega)$, applied to each partial derivative. Possibly, one of the points worth highlighting, from the viewpoint of variational problems, is the fact that functions in $W^{2,p}(\Omega)$ admit traces for functions and all their partial derivatives. In particular, if Ω has a smooth boundary with outer, unit normal $\mathbf{n}$, then

the normal derivative

$$\frac{\partial u}{\partial \mathbf{n}} = \nabla u \cdot \mathbf{n}$$

is well-defined at points in $\partial\Omega$. In particular, we can also talk about the space $W_0^{2,p}(\Omega)$ which is the subspace of $W^{2,p}(\Omega)$ with

$$u = \nabla u = 0 \text{ on } \partial\Omega,$$

or equivalently

$$u = \frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \partial\Omega.$$

From this standpoint, it is important to understand the distinction between the two spaces

$$W_0^{2,p}(\Omega), \quad W_0^{1,p}(\Omega) \cap W^{2,p}(\Omega).$$

Definition 7.6 Let $\Omega \subset \mathbb{R}^N$ be a domain.

1. The subspace

$$W_0^{2,p}(\Omega) \subset W^{2,p}(\Omega)$$

   is the closure in $W^{2,p}(\Omega)$ of the subspace of smooth functions with compact support in Ω.
2. The subspace

$$W_0^{1,p}(\Omega) \cap W^{2,p}(\Omega) \subset W^{2,p}(\Omega)$$

   is the closure in $W^{2,p}(\Omega)$ of the subspace of smooth functions vanishing at $\partial\Omega$.

We will come back to these spaces in the final part of the next chapter, when dealing with second-order variational problems.

7.10 Exercises

1. Let $\Omega \subset \mathbb{R}^N$ be an open subset, and let

$$\{(\boldsymbol{x}_0', t) : t \in J\} \subset \Omega$$

for some fixed $\boldsymbol{x}_0'$ and compact interval $J \subset \mathbb{R}$. Show that if $\phi(x_N)$ is a test function with support contained in J, then there is another test function $\psi(\boldsymbol{x}')$ such that

$$\phi(\boldsymbol{x}) = \psi(\boldsymbol{x}')\phi(x_N)$$

is a test function in Ω with $\psi(\boldsymbol{x}_0') = 1$.

2. Argue that a connected domain Ω in $\mathbb{R}^N$ can be partitioned in disjoints parts Ω_k, $k \in K$, so that over each part Ω_k the decomposition in (7.11) is such that $J_{\boldsymbol{x}'}$ is a true single interval in $\mathbb{R}$.
3. For an arbitrary rotation $\mathbf{R}$ in $\mathbb{R}^N$, $\mathbf{R}^T\mathbf{R} = \mathbf{1}$ ($\mathbf{1}$, the identity matrix), define the function

$$u_{\mathbf{R}}(z) : \mathbf{R}^T\Omega \to \mathbb{R}, \quad u_{\mathbf{R}}(z) = u(\mathbf{R}z),$$

if $\Omega \subset \mathbb{R}^N$ is a domain, and $u \in W^{1,p}(\Omega)$. Prove that $u_{\mathbf{R}} \in W^{1,p}(\mathbf{R}^T\Omega)$, and

$$\nabla u_{\mathbf{R}}(z) = \mathbf{R}^T\nabla u(\mathbf{R}z).$$

4. In the field of PDEs, spaces of functions where not all partial derivatives have the same integrability properties need to be considered. To be specific, consider the space

$$\{u \in L^2(\Omega) : \frac{\partial u}{\partial x_1} \in L^2(\Omega)\}, \quad \boldsymbol{x} = (x_1, x_2), \Omega \subset \mathbb{R}^2,$$

but nothing is required about $\partial u/\partial x_2$. Show that it is a Hilbert space under the inner product

$$\langle u, v\rangle = \int_\Omega [u(\boldsymbol{x})v(\boldsymbol{x}) + \frac{\partial u}{\partial x_1}(\boldsymbol{x})\frac{\partial v}{\partial x_1}(\boldsymbol{x})]\,d\boldsymbol{x}, \quad \boldsymbol{x} = (x_1, x_2).$$

5. Show rigorously that:

   (a) if $u \in W^{1,p}(\Omega)$ and $\psi \in C_c^\infty(\mathbb{R}^N)$, then the product $u\psi$ belongs to $W^{1,p}(\Omega)$, and the product rule holds

   $$\nabla(u\psi) = \nabla u\,\psi + u\nabla\psi \text{ in } \Omega;$$

   (b) if $u \in W_0^{1,p}(\Omega)$ and $\psi \in C_c^\infty(\mathbb{R}^N)$, then the product $u\psi \in W_0^{1,p}(\Omega)$ too.

6. Given a domain $\Omega \subset \mathbb{R}^N$ and a positive volume fraction $t \in (0, 1)$, there is always a sequence of characteristic functions $\{\chi_j(\boldsymbol{x})\}$ of subsets of Ω such that $\chi_j \rightharpoonup t$.
7. Let $\Omega \subset \mathbb{R}^N$ be a bounded, regular domain with a unit, outer normal field $\boldsymbol{n}$ on $\partial\Omega$.

(a) Define the subspace $L^2_{div}(\Omega)$ of fields in $L^2(\Omega; \mathbb{R}^N)$ with a weak divergence in $L^2(\Omega)$.

(b) Consider the further subspace

$$\mathbb{H} = \{\mathbf{F} \in L^2_{div}(\Omega; \mathbb{R}^N) : \operatorname{div} \mathbf{F} = 0\},$$

and check that it is the orthogonal complement of the image, under the gradiente operator, of $H^1_0(\Omega)$.

8. Take $\Omega = (0, 1)^2 \subset \mathbb{R}^2$. Declare a measurable function

$$u(x_1, x_2) : \Omega \to \mathbb{R},$$

as mildly differentiable with respect to x_1, with mild derivative $u_1(x_1, x_2)$, if

$$\lim_{h\to 0} \frac{1}{h^2} \int_\Omega |u(x_1 + h, x_2) - u(x_1, x_2) - hu_1(x_1, x_2)|^2 \, dx_1 \, dx_2 = 0.$$

In a similar way declare a function $u_2(x_1, x_2)$ as the mild derivative of $u(x_1, x_2)$ with respect to x_2. Explore the differences between mild and weak differentiability, starting with smooth functions.

9. Mimic the passage from the rationals to the reals with the absolute value, to define Sobolev spaces from test functions.
10. Show that $W^{1,p}(\mathbb{R}^N) = W^{1,p}_0(\mathbb{R}^N)$.
11. Argue how to define the sequence in Lemma 7.6 from the sequence in (the proof of) Lemma 7.5. Provide the details for the proofs of C^∞-versions of Lemmas 7.5, and 7.6.
12. Redo the proof of Lemma 2.4 in the higher-dimensional setting Corollary 7.1.
13. Let $\mathbf{B} \subset \mathbb{R}^2$ be the unit disc in $\mathbb{R}^2$. Isolate the conditions for functions of a single variable $u = u(x_1)$ to be elements of $H^1(\mathbf{B})$. How would be the situation for a general, regular domain $\Omega \subset \mathbb{R}^2$?
14. With the notation and the ideas of the previous exercise, consider the set of functions

$$\mathbb{L} = \{u(x_1) : u \in W^{1,\infty}(-1, 1)\} \subset H^1(\mathbf{B}).$$

Argue that it is a subspace of $H^1(\mathbf{B})$ that is not closed (under the norm of $H^1(\mathbf{B})$).

Chapter 8
Scalar, Multidimensional Variational Problems

8.1 Preliminaries

Once we have established a solid functional-analytical foundation, we are ready to tackle multidimensional variational problems in which we pretend to minimize the value of the standard, integral functional

$$\int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}$$

among a given set $\mathcal{A}$ of competing functions. The main situation we will explore is that in which feasible functions in $\mathcal{A}$ are determined through their preassigned values around $\partial\Omega$.

We will start with the particular, fundamental case of quadratic functionals which builds upon the Lax-Milgram theorem of Chap. 3. This is rather natural and does not require any new fundamental fact. After that, we will focus on the three important aspects as in Chap. 4, namely,

1. weak lower semicontinuity, the direct method, and one main existence result;
2. optimality conditions in the Hilbert-space scenario, and weak solutions of PDEs;
3. explicit examples.

We will wrap the chapter with a look at the most important example of a second-order problem that is important in applications: the bi-harmonic operator. We will cover, in such a case, much more rapidly the two issues of existence of optimal solutions and optimality.

Since at this point we already have a non-negligible training on most of the abstract, underlying issues, proofs dwell in more technical facts, and, sometimes, they are shorter than the ones for previous similar results.

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4_8

8.2 Abstract, Quadratic Variational Problems

Our abstract discussion on the Lax-Milgram lemma in Sect. 3.2 can be applied directly to quadratic, multi-dimensional variational problems of the form

$$\text{Minimize in } u \in H_0^1(\Omega): \quad \int_\Omega \left[\frac{1}{2} \nabla u(\boldsymbol{x})^T \mathbf{A}(\boldsymbol{x}) \nabla u(\boldsymbol{x}) + f(\boldsymbol{x}) u(\boldsymbol{x}) \right] d\boldsymbol{x}$$

under vanishing Dirichlet boundary conditions around $\partial\Omega$. The important ingredients to ensure the appropriate hypotheses in Theorem 3.1 are

$$\mathbf{A}(\boldsymbol{x}), \text{ symmetric}, \quad |\mathbf{A}(\boldsymbol{x})| \le C, \quad C > 0, \boldsymbol{x} \in \Omega,$$
$$c|\mathbf{u}|^2 \le \mathbf{u}^T \mathbf{A}(\boldsymbol{x}) \mathbf{u}, \quad c > 0, \boldsymbol{x} \in \Omega, \mathbf{u} \in \mathbb{R}^N.$$

In addition $f(\boldsymbol{x}) \in L^2(\Omega)$. Under these assumptions the following is a direct application of Theorem 3.1.

Corollary 8.1 *The variational problem just described admits a unique minimizer $\overline{u} \in H_0^1(\Omega)$ that is characterized by the condition*

$$\int_\Omega [\nabla v(\boldsymbol{x})^T \mathbf{A}(\boldsymbol{x}) \nabla \overline{u}(\boldsymbol{x}) + f(\boldsymbol{x}) v(\boldsymbol{x})] \, d\boldsymbol{x} = 0 \tag{8.1}$$

for every $v \in H_0^1(\Omega)$.

The main specific example is, of course, the Dirichlet principle described in Sect. 1.4, and also mentioned in Sect. 1.8, because of its historical relevance in the development of Functional Analysis and the Calculus of Variations. It corresponds to the choice

$$\mathbf{A}(\boldsymbol{x}) = \mathbf{1}, \quad f(\boldsymbol{x}) \equiv 0,$$

which is obviously covered by Theorem 3.1 and Corollary 8.1. Due to its significance, we state it separately as a corollary.

Corollary 8.2 (Dirichlet's Principle) *For every domain Ω, there is a unique function $u(\boldsymbol{x}) \in H_0^1(\Omega)$ which is a minimizer for the problem*

$$\textit{Minimize in } v(\boldsymbol{x}) \in H^1(\Omega): \quad \frac{1}{2} \int_\Omega |\nabla v(\boldsymbol{x})|^2 \, d\boldsymbol{x}$$

under $u - u_0 \in H_0^1(\Omega)$ for a given $u_0 \in H^1(\Omega)$. This unique function u is determined as a (weak) solution of Laplace's equation

$$\Delta u(\boldsymbol{x}) = 0 \textit{ in } \Omega, \quad u = u_0 \textit{ on } \partial\Omega.$$

As we comment below, much more regularity for the harmonic function u can be derived if one can count on further regularity of the domain Ω, and the boundary datum u_0.

Several variations can be adapted to this model problem. We name a few, some of which are proposed as exercises in the final section. We would have a corresponding corollary for all of them.

- The linear term

$$\int_\Omega f(\boldsymbol{x})u(\boldsymbol{x})\,d\boldsymbol{x}$$

  can be of the form

$$\int_\Omega \mathbf{F}(\boldsymbol{x})\cdot\nabla u(\boldsymbol{x})\,d\boldsymbol{x}$$

  for a field

$$\mathbf{F}(\boldsymbol{x})\in L^2(\Omega;\mathbb{R}^N),$$

  or a combination of both. Even more generally, it could also be

$$\langle f, u\rangle, \quad f\in H^{-1}(\Omega),$$

  if we declare the space $H^{-1}(\Omega)$ as the dual of $H_0^1(\Omega)$. More on this in the final chapter.
- Boundary conditions can be easily changed to a non-vanishing situation, as in the Dirichlet's principle, by selecting some appropriate specific function $u_0 \in H^1(\Omega)$, and defining the bilinear form over functions of the kind $u + u_0$ for $u \in H_0^1(\Omega)$.
- The bilinear form, i.e. the integral functional, can also incorporate a quadratic term in u

$$\int_\Omega \left[\frac{1}{2}\nabla u(\boldsymbol{x})^T\mathbf{A}(\boldsymbol{x})\nabla u(\boldsymbol{x}) + \frac{1}{2}a(\boldsymbol{x})u(\boldsymbol{x})^2 + f(\boldsymbol{x})u(\boldsymbol{x})\right] d\boldsymbol{x}$$

  for a non-negative coefficient $a(\boldsymbol{x})$.
- One can also consider the same variational problems without an explicit boundary condition like

$$\text{Minimize in } u\in H^1(\Omega):$$
$$\int_\Omega \left[\frac{1}{2}\nabla u(\boldsymbol{x})^T\mathbf{A}(\boldsymbol{x})\nabla u(\boldsymbol{x}) + \frac{1}{2}u(\boldsymbol{x})^2 + f(\boldsymbol{x})u(\boldsymbol{x})\right] d\boldsymbol{x}.$$

The minimization process in this situation leads to the so-called natural boundary condition, as it is found as a result of the optimization process itself, and not imposed in any way.

- Boundary conditions may come in the Neumann form for the normal derivative

$$\text{Minimize in } u \in H^1(\Omega): \\ \int_\Omega \left[\frac{1}{2}\nabla u(\boldsymbol{x})^T \mathbf{A}(\boldsymbol{x})\nabla u(\boldsymbol{x}) + \frac{1}{2}u(\boldsymbol{x})^2 + f(\boldsymbol{x})u(\boldsymbol{x})\right] d\boldsymbol{x}.$$

  under

$$\nabla u(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x}) = h(\boldsymbol{x}) \text{ on } \partial\Omega,$$

  for a function h defined on $\partial\Omega$, and $\boldsymbol{n}(\boldsymbol{x})$, the unit, outer normal to $\partial\Omega$. Ω is supposed to have a boundary which is a smooth $N-1$ compact hypersurface. This problem would require, however, a much more detailed analysis to determine Lebesgue spaces defined on manifolds for functions h.
- The matrix field $\mathbf{A}(\boldsymbol{x})$ might not be symmetric.

There are two important issues to explore, once one has shown existence of optimal solutions for a certain variational principle.

1. In the first place, it is most crucial to establish optimality conditions. These are additional fundamental requirements that optimal solutions must comply with precisely because they are optimal solutions for a given variational problem; they typically involved, in the multidimensional case, partial differential equations (PDEs). These are initially formulated in a weak form; in the case of quadratic functionals, such optimality conditions are given as a fundamental part of the application of the Lax-Milgram lemma. In Corollary 8.1, it is (8.1). We say that (8.1) is the weak form of the linear, elliptic PDE

$$-\operatorname{div}[\mathbf{A}(\boldsymbol{x})\nabla\overline{u}(\boldsymbol{x})] + f(\boldsymbol{x}) = 0 \text{ in } \Omega, \quad \overline{u} = 0 \text{ on } \partial\Omega,$$

   because by multiplying formally this equation by an arbitrary function $v \in H_0^1(\Omega)$ and integrating by parts, we recover (8.1).
2. The second important issue is to derive further regularity properties (better integrability, continuity, smoothness, analyticity, etc) for the optimal solution $\overline{u}$ from additional conditions on the ingredients of the problem: the domain Ω, and $\mathbf{A}(\boldsymbol{x})$ and $f(\boldsymbol{x})$. Typically, this regularity issue is delicate and requires quite a bit of fine work.

We will be using these quadratic model problem for a more general integral functional. However, when one is off the quadratic framework, the existence of optimal solution is much more involved as there is no general result as neat as the Lax-Milgram lemma.

8.3 Scalar, Multidimensional Variational Problems

We are now ready to tackle scalar, multi-dimensional variational problems of the form

$$I(u) = \int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x} \tag{8.2}$$

under typical constraints prescribing boundary values

$$u = u_0 \text{ on } \partial\Omega \tag{8.3}$$

for a fixed, given function u_0. Since such variational problems will be setup in Sobolev spaces $W^{1,p}(\Omega)$ for exponent $p > 1$, the underlying set Ω will always be taken to be a bounded domain according to Definition 7.2. Moreover the function u_0 will belong to the same space $W^{1,p}(\Omega)$ and competing functions $u \in W^{1,p}(\Omega)$ will be further restricted by putting

$$u - u_0 \in W_0^{1,p}(\Omega).$$

This is our formal way to enforce (8.3). More specifically, we will be concerned with the variational problem

$$\text{Minimize in } u \in \mathcal{A}: \quad I(u) = \int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}$$

where $\mathcal{A} \subset W^{1,p}(\Omega)$ is a non-empty, weakly-closed subset. We would like to learn what structural conditions on the integrand

$$F(\boldsymbol{x}, u, \mathbf{u}) : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}, \tag{8.4}$$

and on such feasible set $\mathcal{A}$ of competing functions, guarantee the success of the direct method Proposition 3.1. In this section we focus on the weak lower semicontinuity property. This will be a direct corollary of the following general result.

Theorem 8.1 *Let*

$$F(\mathbf{u}, \mathbf{v}) : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$$

be continuous, and bounded from below. Consider the associated integral functional

$$E(\mathbf{u}, \mathbf{v}) = \int_\Omega F(\mathbf{u}(\boldsymbol{x}), \mathbf{v}(\boldsymbol{x}))\, d\boldsymbol{x}$$

for pairs

$$(\mathbf{u}, \mathbf{v}) \in L^p(\Omega; \mathbb{R}^m \times \mathbb{R}^n).$$

Then E is strong-weak lower semicontinuous, i.e.

$$E(\mathbf{u}, \mathbf{v}) \leq \liminf_{j\to\infty} E(\mathbf{u}_j, \mathbf{v}_j)$$

whenever

$$\mathbf{u}_j \to \mathbf{u} \text{ in } L^p(\Omega; \mathbb{R}^m), \quad \mathbf{v}_j \rightharpoonup \mathbf{v} \text{ in } L^p(\Omega; \mathbb{R}^n), \tag{8.5}$$

if and only if $F(\mathbf{u}, \cdot)$ is convex for every $\mathbf{u} \in \mathbb{R}^m$.

Proof We follow along the lines of the similar proof for dimension 1 Theorem 4.2. Suppose we have the convergence (8.5). For each j, by Jensen's inequality we have

$$F\left(\tilde{\mathbf{u}}_j, \frac{1}{|\tilde{\Omega}|}\int_{\tilde{\Omega}} \mathbf{v}_j(\boldsymbol{x})\,d\boldsymbol{x}\right) \leq \frac{1}{|\tilde{\Omega}|}\int_{\tilde{\Omega}} F(\tilde{\mathbf{u}}_j, \mathbf{v}_j(\boldsymbol{x}))\,d\boldsymbol{x}$$

for arbitrary subsets $\tilde{\Omega} \subset \Omega$, and

$$\tilde{\mathbf{u}}_j = \frac{1}{|\tilde{\Omega}|}\int_{\tilde{\Omega}} \mathbf{u}_j(\boldsymbol{x})\,d\boldsymbol{x}.$$

In particular, if $\{\Omega_i\}_i$ is a finite, arbitrary partition of Ω, and

$$\tilde{\mathbf{u}}_j^{(i)} = \frac{1}{|\Omega_i|}\int_{\Omega_i} \mathbf{u}_j(\boldsymbol{x})\,d\boldsymbol{x},$$

we will have

$$\begin{aligned}\sum_i |\Omega_i| F\left(\tilde{\mathbf{u}}_j^{(i)}, \frac{1}{|\Omega_i|}\int_{\Omega_i} \mathbf{v}_j(\boldsymbol{x})\,d\boldsymbol{x}\right) &\leq \sum_i \int_{\Omega_i} F(\tilde{\mathbf{u}}_j^{(i)}, \mathbf{v}_j(\boldsymbol{x}))\,d\boldsymbol{x} \\ &\leq \int_{\Omega} F(\mathbf{u}_j(\boldsymbol{x}), \mathbf{v}_j(\boldsymbol{x}))\,d\boldsymbol{x} + R_j,\end{aligned}$$

where

$$R_j = \sum_i R_{j,i}, \quad R_{j,i} = \int_{\Omega_i} |F(\mathbf{u}_j(\boldsymbol{x}), \mathbf{v}_j(\boldsymbol{x})) - F(\tilde{\mathbf{u}}_j^{(i)}, \mathbf{v}_j(\boldsymbol{x}))|\,d\boldsymbol{x}.$$

If we select the partition $\{\Omega_i\}_i$ of Ω, depending on j, in such a way that

$$\mathbf{u}_j(\boldsymbol{x}) - \sum_i \chi_{\Omega_i}(\boldsymbol{x})\tilde{\mathbf{u}}_j^{(i)} \to 0$$

pointwise in Ω, the sequence of functions

$$r_j(\boldsymbol{x}) = \sum_i \chi_{\Omega_i}(\boldsymbol{x})|F(\mathbf{u}_j(\boldsymbol{x}), \mathbf{v}_j(\boldsymbol{x})) - F(\tilde{\mathbf{u}}_j^{(i)}, \mathbf{v}_j(\boldsymbol{x}))|, \quad R_j = \int_\Omega r_j(\boldsymbol{x})\,d\boldsymbol{x},$$

converges pointwise to zero in Ω. Let $E \subset \Omega$ be a measurable subset in which the sequence of remainders $\{r_j\}$ is uniformly bounded. If all the previous integrals are restricted to E, by the Lebesgue dominated convergence theorem, we would find after taking limits in j, because

$$R_j(E) \equiv \int_E r_j(\boldsymbol{x})\,d\boldsymbol{x} \to 0,$$

that

$$\int_E F(\mathbf{u}(\mathbf{x}), \mathbf{v}(\mathbf{x}))\,d\boldsymbol{x} \le \liminf_{j\to\infty} \int_E F(\mathbf{u}_j(\boldsymbol{x}), \mathbf{v}_j(\boldsymbol{x}))\,d\boldsymbol{x}.$$

Since F can be assumed, without loss of generality to be non-negative $F \ge 0$, then

$$\int_E F(\mathbf{u}(\mathbf{x}), \mathbf{v}(\mathbf{x}))\,d\boldsymbol{x} \le \liminf_{j\to\infty} \int_\Omega F(\mathbf{u}_j(\boldsymbol{x}), \mathbf{v}_j(\boldsymbol{x}))\,d\boldsymbol{x},$$

and the arbitrariness of E filling out all of Ω, leads to the sufficiency.

For the necessity, we invoke a generalization of Example 2.12 in the higher dimensional setting. Given a domain $\Omega \subset \mathbb{R}^N$ and a positive volume fraction $t \in (0, 1)$, there is always a sequence of characteristic functions $\{\chi_j(\boldsymbol{x})\}$ of subsets of Ω such that $\chi_j \rightharpoonup t|\Omega|$. This is left as an exercise. If we now consider the sequence of pairs

$$(\mathbf{u}, \chi_j(\boldsymbol{x})\mathbf{v}_1 + (1 - \chi_j(\boldsymbol{x}))\mathbf{v}_0),$$

for vectors $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{v}_1, \mathbf{v}_0 \in \mathbb{R}^n$, we see that

$$\chi_j(\boldsymbol{x})\mathbf{v}_1 + (1 - \chi_j(\boldsymbol{x}))\mathbf{v}_0 \rightharpoonup t\mathbf{v}_1 + (1 - t)\mathbf{v}_0,$$

and by the strong-weak lower semicontinuity we would conclude that

$$\begin{aligned}
|\Omega| F(\mathbf{u}, t\mathbf{v}_1 + (1-t)\mathbf{v}_0) \leq & \liminf_{j\to\infty} \int_\Omega F(\mathbf{u}, \chi_j(\boldsymbol{x})\mathbf{v}_1 + (1-\chi_j(\boldsymbol{x}))\mathbf{v}_0)\, d\boldsymbol{x} \\
= & \liminf_{j\to\infty} \int_\Omega \left[\chi_j(\boldsymbol{x}) F(\mathbf{u}, \mathbf{v}_1) + (1-\chi_j(\boldsymbol{x})) F(\mathbf{u}, \mathbf{v}_0)\right] d\boldsymbol{x} \\
= & |\Omega| (t F(\mathbf{u}, \mathbf{v}_1) + (1-t) F(\mathbf{u}, \mathbf{v}_0)).
\end{aligned}$$

□

Note how this proof has turned out much simpler than its one-dimensional twin Theorem 4.2 because in the current higher-dimensional situations we have decoupled the two sets of variables **u** and **v**, i.e. they are unrelated.

Corollary 8.3 *Suppose the integrand in* (8.4) *is measurable in x, and continuous in pairs* $(u, \mathbf{u})$ *and bounded from below. Then functional* (8.2) *is weakly lower semicontinuous in* $W^{1,p}(\Omega)$ *if and only if F is convex in* **u** *for a.e.* $x \in \Omega$, *and* $u \in \mathbb{R}$.

Proof The sufficiency part of this corollary is a direct consequence of the previous theorem if we identify

$$\mathbf{U}_j(\boldsymbol{x}) = (\boldsymbol{x}, \mathbf{u}_j(\boldsymbol{x}))$$

in such a way that the strong convergence of $\{\mathbf{U}_j\}$ in $L^p(\Omega; \mathbb{R}^{N+m})$ is equivalent to the convergence of $\{\mathbf{u}_j\}$ in $L^p(\Omega; \mathbb{R}^m)$. The necessity is a bit more involved than the property shown in Theorem 8.1 precisely by the remark made right after the proof of it. But it is a clear indication that it should be correct; and indeed it is. Its proof would require finer calculations as in the proof of Theorem 4.2. We do not insist on this point because our main interest is in the sufficiency part. □

8.4 A Main Existence Theorem

Corollary 8.3 is the cornerstone of a general existence theorem for optimal solutions for variational problems with an integral functional of type (8.2)

$$I(u) = \int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}$$

under boundary values around $\partial\Omega$ for admissible u's, and additional assumptions ensuring coercivity in appropriate Sobolev spaces.

Theorem 8.2 *Let*

$$F(\boldsymbol{x}, u, \mathbf{u}) : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$$

be measurable with respect to $\boldsymbol{x}$*, and continuous with respect to pairs* $(u, \mathbf{u})$*. Suppose in addition that.*

1. *there is* $p > 1$ *and* $C > 0$ *with*

$$C(|\mathbf{u}|^p - 1) \le F(\boldsymbol{x}, u, \mathbf{u}) \tag{8.6}$$

   for every triplet $(\boldsymbol{x}, u, \mathbf{u})$*;*
2. *the dependence*

$$\mathbf{u} \mapsto F(\boldsymbol{x}, u, \mathbf{u})$$

   is convex for every pair $(\boldsymbol{x}, u)$.

Then for every $u_0 \in W^{1,p}(\Omega)$*, furnishing boundary values, the variational problem*

$$\textit{Minimize in } u \in W^{1,p}(\Omega): \quad I(u)$$

under

$$u - u_0 \in W_0^{1,p}(\Omega)$$

admits optimal solutions.

Proof The proof reduces to the use of the direct method Proposition 3.1 together with Corollary 8.3 for the weak lower semicontinuity.

- Note how the coercivity condition (8.6), in addition to the Dirichlet boundary condition $u - u_0 \in W_0^{1,p}(\Omega)$ imposed on competing functions, imply that minimizing sequences are uniformly bounded in $W^{1,p}(\Omega)$. In fact, if $\{u_0 + u_j\}$ is minimizing with

$$\{u_j\} \subset W_0^{1,p}(\Omega),$$

  then

$$\|\nabla u_0 + \nabla u_j\|_{L^p(\Omega)}^p \le 1 + \frac{1}{C} I(u_0 + u_j),$$

  if I is the corresponding integral functional. Thus, if $\{u_0+u_j\}$ is truly minimizing with a value $m \in \mathbb{R}$ for the infimum,

$$\|\nabla u_0 + \nabla u_j\|_{L^p(\Omega)} \le \left(1 + \frac{m+1}{C}\right)^{1/p},$$

and

$$\begin{aligned}\|\nabla u_j\|_{L^p(\Omega)} \leq &\|\nabla u_0 + \nabla u_j\|_{L^p(\Omega)} + \|\nabla u_0\|_{L^p(\Omega)}\\ \leq &\left(1 + \frac{m+1}{C}\right)^{1/p} + \|\nabla u_0\|_{L^p(\Omega)},\end{aligned}$$

for all j large, and a constant on the right-hand side independent of j.

- Because $p > 1$, there are subsequences converging weakly in $W^{1,p}(\Omega)$ and strongly in $L^p(\Omega)$ to limit functions u that, therefore, comply with the same boundary condition, and are hence admissible.
- The convexity of the integrand with respect to the gradient variable yields, according to Corollary 8.3, a functional I which is weakly lower semicontinuity in $W^{1,p}(\Omega)$.

We therefore are in possession of all required ingredients for the direct method Proposition 3.1 to work, and conclude the existence of minimizers. □

Boundary conditions allow for some variations. Some of these are proposed in the exercises, as well as specific problems for concrete integrands. One important such situation is that of the natural boundary condition (Exercise 1). Another important, even more general situation corresponds to a feasible set $\mathcal{A}$ which is weakly closed in the appropriate space. Such a subset $\mathcal{A}$ must be weakly closed to ensure that weak limits of sequences in $\mathcal{A}$ stay within $\mathcal{A}$. A paradigmatic situation corresponds to the classical obstacle problem Exercise 29. We will see some other relevant situation in the next chapter.

Uniqueness of optimal solutions require, as usual, more demanding conditions on the integrand: jointly strict convexity with respect to $(u, \mathbf{u})$. In general, it is not sufficient to have strict convexity only with respect to $\mathbf{u}$. As a matter of fact, lack of strict or even plain convexity with respect to u, having the same strictly convex dependence on $\mathbf{u}$, yields a lot a flexibility to study interesting problems.

Proposition 8.1 *Suppose, in addition to hypotheses in Theorem 8.2, that the integrand F is jointly strictly convex with respect to pairs $(u, \mathbf{u})$. Then there is a unique optimal solution of the corresponding variational problem under given Dirichlet boundary conditions determined by $u_0 \in W^{1,p}(\Omega)$.*

The proof is similar to that of Proposition 3.3.

8.5 Optimality Conditions: Weak Solutions for PDEs

We now come to the last of the main issues concerning variational problems, that of deriving optimality conditions that optimal solutions of problems must comply with precisely because they are optimal. As usual, the natural scenario to derive optimality is that of Hilbert spaces, though it can be generalized to Banach spaces.

For this reason, we restrict attention in this section to the case $p = 2$, and so we will be working in the space $H^1(\Omega)$. For the sake of simplicity, we avoid the irrelevant dependence of the integrand $F(\boldsymbol{x}, u, \mathbf{u})$ on $\boldsymbol{x}$. We are not writing the most general conditions possible, but will be contented with understanding the nature of such optimality conditions. Note that no convexity is assumed on the next statement of optimality conditions. It is a necessary statement.

Theorem 8.3 *Suppose the integrand*

$$F(u, \mathbf{u}) : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$$

is continuously differentiable with partial derivatives

$$F_u(u, \mathbf{u}), \quad F_{\mathbf{u}}(u, \mathbf{u}),$$

and, for some constant $C > 0$,

$$|F(u, \mathbf{u})| \le C(|\mathbf{u}|^2 + |u|^2),$$

$$|F_u(u, \mathbf{u})| \le C(|\mathbf{u}| + |u|), \quad |F_{\mathbf{u}}(u, \mathbf{u})| \le C(|\mathbf{u}| + |u|).$$

If a function $v \in H^1(\Omega)$ is a minimizer for the problem

$$\textit{Minimize in } u \in H^1(\Omega) : \quad I(u) = \int_\Omega F(u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}$$

under

$$u - u_0 \in H_0^1(\Omega), \quad u_0 \in H^1(\Omega), \textit{ given},$$

then

$$\int_\Omega [F_u(v(\boldsymbol{x}), \nabla v(\boldsymbol{x}))V(\boldsymbol{x}) + F_{\mathbf{u}}(v(\boldsymbol{x}), \nabla v(\boldsymbol{x})) \cdot \nabla V(\boldsymbol{x})]\, d\boldsymbol{x} = 0 \tag{8.7}$$

for every $V \in H_0^1(\Omega)$.

Condition (8.7) is the weak form of the PDE which is known as the Euler-Lagrange equation of the problem

$$-\operatorname{div}[F_{\mathbf{u}}(\boldsymbol{x}, v(\boldsymbol{x}), \nabla v(\boldsymbol{x})] + F_u(\boldsymbol{x}, v(\boldsymbol{x}), \nabla v(\boldsymbol{x}) = 0 \text{ in } \Omega, \quad v = v_0 \text{ on } \partial\Omega.$$

Note how (8.7) is precisely what you find when, formally, multiply in this differential equation by an arbitrary $V \in H_0^1(\Omega)$, and perform an integration by parts, as usual. Solutions v of (8.7) are called critical functions (of the corresponding functional).

Proof Note that under the bounds assumed on the sizes of $F(u, \mathbf{u})$ and its partial derivatives $F_u(u, \mathbf{u})$ and $F_{\mathbf{u}}(u, \mathbf{u})$, the integral in (8.7) is well-defined for every $V \in H_0^1(\Omega)$.

The proof of (8.7) is very classical. It follows the basic strategy in the corresponding abstract result Proposition 3.4. Let $v \in H_0^1(\Omega)$ be a true minimizer for the variational problem in the statement, and let $V \in H_0^1(\Omega)$ be an arbitrary element. For every real $s \in \mathbb{R}$, the section

$$g(s) = E(v + sV)$$

is a function with a global minimum at $s = 0$. If g is differentiable, then we should have $g'(0) = 0$. We have

$$g(s) = \int_\Omega F(v(\boldsymbol{x}) + sV(\boldsymbol{x}), \nabla v(\boldsymbol{x}) + s\nabla V(\boldsymbol{x}))\, d\boldsymbol{x}.$$

Since the integrand $F(u, \mathbf{u})$ is continuously differentiable, and the result of the formal differentiation with respect to s

$$\int_\Omega F_u(v(\boldsymbol{x}) + sV(\boldsymbol{x}), \nabla v(\boldsymbol{x}) + s\nabla V(\boldsymbol{x}))V(\boldsymbol{x})$$
$$+ F_{\mathbf{u}}(v(\boldsymbol{x}) + sV(\boldsymbol{x}), \nabla v(\boldsymbol{x}) + s\nabla V(\boldsymbol{x})) : \nabla V(\boldsymbol{x})\, d\boldsymbol{x}$$

is well-defined because of our upper bounds on the size of F and its partial derivatives, it is easy to argue that indeed g is differentiable, and its derivative is given by the last integral. The condition for global minimum $g'(0) = 0$ then becomes exactly (8.7). □

A final fundamental issue is that of the sufficiency of optimality conditions. Suppose we have a function $v \in H^1(\Omega)$ for which (8.7) holds for every $V \in H_0^1(\Omega)$. Under what conditions on the integrand $F(u, \mathbf{u})$ can we be sure that v is in fact a minimizer of the corresponding variational problem? This again involves convexity. One main point of research, that we will barely touch upon in the next chapter, is to derive results of existence of critical functions which are not minimizers for different families of integral functionals.

Theorem 8.4 *Suppose the integrand*

$$F(u, \mathbf{u}) : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$$

is like has been indicated in Theorem 8.3, and it is jointly convex in pairs $(u, \mathbf{u})$. *Suppose that a certain function* $v \in H^1(\Omega)$, *with* $v - u_0 \in H_0^1(\Omega)$ *and* u_0 *prescribed, is such that* (8.7) *holds for every* $V \in H_0^1(\Omega)$. *Then* v *is a minimizer for the problem*

$$\textit{Minimize in } u \in H^1(\Omega) : \quad I(u) = \int_\Omega F(u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}$$

under

$$u - u_0 \in H_0^1(\Omega).$$

Proof The proof is essentially the one for Theorem 4.5 adapted to a multidimensional situation. Suppose the feasible function v is such that (8.7) holds for every $V \in H_0^1(\Omega)$. The sum $v+V$ is also feasible for the variational problem, and because of the convexity assumed on F, for a.e. $\boldsymbol{x} \in \Omega$,

$$\begin{aligned} F(v(\boldsymbol{x}) + V(\boldsymbol{x}), \nabla v(\boldsymbol{x}) + \nabla V(\boldsymbol{x})) \geq & F(v(\boldsymbol{x}), \nabla v(\boldsymbol{x})) \\ & + F_u(v(\boldsymbol{x}), \nabla v(\boldsymbol{x}))V(\boldsymbol{x}) \\ & + F_{\mathbf{u}}(v(\boldsymbol{x}), \nabla v(\boldsymbol{x})) : \nabla V(\boldsymbol{x}). \end{aligned}$$

Upon integration on $\boldsymbol{x}$, taking into account (8.7), we see that $I(v + V) \geq I(v)$, for arbitrary $V \in H_0^1(\Omega)$, and v is indeed a minimizer of the problem. □

8.6 Variational Problems in Action

We have seen that the two main ingredients that ensure existence of optimal solutions for a typical variational problem associated with an integral functional

$$\int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}, \quad F(\boldsymbol{x}, u, \mathbf{u}) : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$$

is the convexity and coercivity of the integrand $F(\boldsymbol{x}, u, \mathbf{u})$ with respect to the gradient variable $\mathbf{u}$. Though readers may think that once understood this, one can be reputed to know a lot about variational problems, the truth is that the application of Theorem 8.2 to concrete examples might turned out to be more challenging than anticipated.

One needs to deal with convexity in the most efficient way, and this implies to know the main operations among functions that preserve convexity in order to have a clear picture of how to build more sophisticated convex functions from simple ones. We assume that readers are familiar with the use of the positivity of the hessian to check convexity for smooth cases. This is taught in Multivariate Calculus courses. Often times, however, the application of such criterium, though clear as a procedure, may be not so easy to implement. For instance, the function

$$F(\mathbf{u}) = \frac{1}{2}|\mathbf{u}|^2 + |\mathbf{u}|\mathbf{u} \cdot \mathbf{v}, \quad \mathbf{v} \in \mathbb{R}^N, |\mathbf{v}| \leq \frac{1}{2},$$

is convex, but it is not so straightforward to check it. Same comments apply to coercivity. At times, convexity may be quite straightforward, and it is coercivity

that turns out to be more involved. We aim at giving some practical hints that may help in dealing with examples.

Recall again (Definition 4.1) that a function

$$\phi : D \subset \mathbb{R}^N \to \mathbb{R}$$

is convex if D is a convex set, and

$$\phi(t_1\mathbf{x}_1 + t_2\mathbf{x}_2) \le t_1\phi(\mathbf{x}_1) + t_2\phi(\mathbf{x}_2), \quad \mathbf{x}_1, \mathbf{x}_2 \in D, t_1, t_2 \ge 0, t_1 + t_2 = 1.$$

If there is no mention about the set D, one consider it to be the whole space $\mathbb{R}^N$ or the natural domain of definition of ϕ (which must be a convex set).

The following is elementary, but they are the principles to build new convex functions from old ones. We have already mentioned it in Sect. 4.1.

Proposition 8.2

1. *Every linear (affine) function is convex (and concave).*
2. *A linear combination of convex functions with positive scalars is convex. In particular, the sum of convex functions is convex, and so is the product by a positive number.*
3. *The composition of a convex function with an increasing, convex function of one variable is convex.*
4. *The supremum of convex functions is convex.*

We already know (Proposition 4.1) that, when ϕ is convex, then

$$\mathbf{C}_L\phi = \sup\{\psi : \psi \le \phi, \psi, \text{ linear}\} \equiv \phi.$$

If it is not, we isolate the following important concept.

Definition 8.1 For a function $\phi(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}$, we define the convex function

$$\mathbf{C}_L\phi \equiv \mathbf{C}_C\phi \equiv \mathbf{C}\phi$$

as its convexification. It is the largest convex function, not bigger than ϕ itself.

Concerning coercivity, we just want to mention a main principle that plays an important role in many different contexts when contributions in inequalities are desired to be absorbed by other terms. It is just a simple generalization of inequality (4).

Lemma 8.1 *For* $\epsilon > 0$, *and* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$,

$$\mathbf{u} \cdot \mathbf{v} \le \frac{\epsilon^2}{2}|\mathbf{u}|^2 + \frac{1}{2\epsilon^2}|\mathbf{v}|^2.$$

Some examples follow.

Example 8.1 The size of a vector $\mathbf{u} \mapsto |\mathbf{u}|$ is a convex function. Indeed, because

$$|\mathbf{u}| = \sup_{|\mathbf{v}|=1} \mathbf{v} \cdot \mathbf{u},$$

it is the supremum of linear functions. Hence, all functions of the form $\phi(\mathbf{u}) = f(|\mathbf{u}|)$ for a non-decreasing, convex function f of a single variable, are convex. In particular,

$$|\mathbf{u}|^p,\ p \geq 1, \quad \sqrt{1 + |\mathbf{u}|^q},\ q \geq 2.$$

Example 8.2 If $\psi(\mathbf{u})$ is convex, then $\phi(\mathbf{u}) = \psi(\mathbf{A}\mathbf{u})$ is convex as well for every matrix $\mathbf{A}$. In this way the functions

$$|\mathbf{A}\mathbf{u}|^p,\ p \geq 1, \quad \sqrt{1 + |\mathbf{A}\mathbf{u}|^q},\ q \geq 2,$$

are convex too.

Example 8.3 This is the one written above

$$F(\mathbf{u}) = \frac{1}{2}|\mathbf{u}|^2 + |\mathbf{u}|\mathbf{u} \cdot \mathbf{v}, \quad \mathbf{v} \in \mathbb{R}^N,\ |\mathbf{v}| \leq \frac{1}{2}. \tag{8.8}$$

One first realizes that for an additional constant vector $\mathbf{w} \in \mathbb{R}^N$, we have that the quadratic function

$$\mathbf{u} \mapsto \frac{1}{2}|\mathbf{u}|^2 + \mathbf{u} \cdot \mathbf{w}\, \mathbf{u} \cdot \mathbf{v} = \mathbf{u} \left(\frac{1}{2}\mathbf{1} + \mathbf{w} \otimes \mathbf{v} \right) \mathbf{u}$$

is convex (because it is positive definite) provided $|\mathbf{w}| \leq 1$, $|\mathbf{v}| \leq 1/2$. Hence if we take the supremum in $\mathbf{w}$, the resulting function is convex too, i.e. $F(\mathbf{u})$ in (8.8) is convex.

8.7 Some Examples

Quadratic variational problems are associated with linear Euler-Lagrange equations, and can be tackled within the scope of the Lax-Milgram lemma as we have already seen. In this section, we would like to mention some non-linear problems that come directly from our basic results for scalar, multidimensional variational problems. Some examples are based on the ones in the last section.

1. One first easy example is the one corresponding to the integrand

$$F(\boldsymbol{x}, u, \mathbf{u}) = \frac{1}{2}|\mathbf{u}|^2 + |\mathbf{u}|\, \mathbf{u} \cdot \mathbf{F}(\boldsymbol{x})$$

for a given field $\mathbf{F}(\boldsymbol{x})$ such that

$$|\mathbf{F}(\boldsymbol{x})| \leq 1/2, \quad \boldsymbol{x} \in \Omega.$$

It admits the variant

$$F(\boldsymbol{x}, u, \mathbf{u}) = \frac{1}{2}|\mathbf{u}|^2 + \sqrt{1 + |\mathbf{u}|^2}\,\mathbf{u} \cdot \mathbf{F}(\boldsymbol{x}).$$

The existence of minimizers under standard Dirichlet boundary conditions is straightforward. Coercivity can be treated through Lemma 8.1. The associated Euler-Lagrange equations are a bit intimidating. For the first case, it can be written in the form

$$\operatorname{div}\left[\nabla u + |\nabla u|\left(\mathbf{1} + \frac{\nabla u}{|\nabla u|} \otimes \frac{\nabla u}{|\nabla u|}\right)\mathbf{F}\right] = 0,$$

where the rank-one matrix $\mathbf{u} \otimes \mathbf{v}$, for two vectors $\mathbf{u}$, $\mathbf{v}$ is given by

$$(\mathbf{u} \otimes \mathbf{v})_{ij} = u_i v_j.$$

Note that the existence of a unique minimizer for the integral functional with one of these two integrands immediately yields the existence of (weak) solutions of, for example, this last non-linear PDE.

2. Our second example is of the form

$$F(\boldsymbol{x}, u, \mathbf{u}) = \frac{1}{2}\left(|\mathbf{u}| + f(\boldsymbol{x})\right)^2 + u(\boldsymbol{x})g(\boldsymbol{x}),$$

with $f \in L^1(\Omega)$, and $g \in L^2(\Omega)$. Once again, the existence of a unique minimizer does not require any special argument, and such minimizer is, therefore, a (weak) solution of the non-linear PDE

$$\operatorname{div}\left(\nabla u + \frac{\nabla u}{|\nabla u|} f\right) = g$$

under typical boundary conditions.

3. Consider

$$F(\boldsymbol{x}, u, \mathbf{u}) = \frac{1}{2}|\mathbf{u}|^2 + f(u)$$

where $f(u)$ is a real, non-negative, continuous function. The existence of minimizers is straightforward. Without further conditions on the function f, it

is not possible to guarantee uniqueness. Compare the two cases

$$f(u)=\frac{1}{2}u^2,\quad f(u)=\frac{1}{2}(|u|-1)^2.$$

4. For exponent p greater than 1, but different from 2, the parallel functional for the pth-Dirichlet problem is

$$\text{Minimize in } u(\boldsymbol{x})\in W^{1,p}(\Omega):\quad \int_\Omega \frac{1}{p}|\nabla u(\boldsymbol{x})|^p\,d\boldsymbol{x}$$

under prescribed boundary conditions $u-u_0\in W_0^{1,p}(\Omega)$ for a given $u_0\in W^{1,p}(\Omega)$. The associated Euler-Lagrange equation reads

$$\operatorname{div}(|\nabla u|^{p-2}\nabla u)=0 \text{ in } \Omega.$$

This non-linear PDE is known as the pth-laplacian equation for obvious reasons, and it probably is the best studied one after Laplaces'. Even for the case $p=1$, a lot of things are known though the analysis is much more delicate. The existence of a unique minimizer for the underlying variational problem is, however, a direct consequence of our analysis in this chapter.
5. The functional for non-parametric minimal surfaces

$$\int_\Omega \sqrt{1+|\nabla u(\boldsymbol{x})|^2}\,d\boldsymbol{x}$$

faces the same delicate issue of the linear growth so that $W^{1,1}(\Omega)$ is the natural space. It is also a problem that requires a lot of expertise. One can always try a small regularization of the form

$$\int_\Omega \left(\frac{\epsilon}{2}|\nabla u|^2+\sqrt{1+|\nabla u(\boldsymbol{x})|^2}\right)d\boldsymbol{x}$$

for some small, positive parameter ϵ. Or even better

$$F_\epsilon(\mathbf{u})=\begin{cases}\sqrt{1+|\mathbf{u}|^2}, & |\mathbf{u}|\le\epsilon^{-1},\\ \epsilon|\mathbf{u}|^2+((1+\epsilon^2)^{-1/2}-2)|\mathbf{u}|+\epsilon(1+\epsilon^2)^{-1/2}+\epsilon^{-1}, & |\mathbf{u}|\ge\epsilon^{-1}.\end{cases}$$

The form of this integrand has been adjusted for the zone $|\mathbf{u}|\ge\epsilon^{-1}$ in such way that it has quadratic growth, and the overall resulting integrand F_ϵ turns out to be at least C^1. The variational problem for F_ϵ has a unique minimizer u_ϵ. This is the easy part. The whole point of such regularization is to investigate the behavior of such minimizers as ϵ tends to zero. This is typically reserved for a much more specialized analysis.

6. Our final examples are of the form

$$F(u, \mathbf{u}) = \frac{a(u)}{2}|\mathbf{u}|^2, \quad F(u, \mathbf{u}) = \mathbf{u}^T \mathbf{A}(u)\mathbf{u},$$

where the respective matrices involved

$$\mathbf{A}(u) = \frac{1}{2}a(u)\mathbf{1}, \quad \mathbf{A}(u),$$

need to be, regardless of its dependence on u, uniformly bounded and positive definite. The existence of minimizers can then be derived directly from our results.

8.8 Higher-Order Variational Principles

Once we have spaces $W^{2,p}(\Omega)$ to our disposal we can treat variational problems of second-order, whose model problem is of the form

$$\text{Minimize in } u \in W^{2,p}(\Omega): \quad I(u) = \int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}), \nabla^2 u(\boldsymbol{x}))\, d\boldsymbol{x},$$

typically, under boundary conditions on competing functions that might involve u and/or ∇u around $\partial\Omega$. In full generality, the density

$$F(\boldsymbol{x}, u, \mathbf{u}, \mathbf{U}) : \Omega \times \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^{N\times N} \to \mathbb{R} \tag{8.9}$$

is supposed to be, at least, continuous in triplets $(u, \mathbf{u}, \mathbf{U})$ and measurable in $\boldsymbol{x} \in \Omega$. $\Omega \subset \mathbb{R}^N$ is a smooth, regular, bounded domain.

We could certainly cover similar facts for second-order problems as the ones stated and proved in Sects. 8.3, 8.4, and 8.5 for first-order problem, but we trust, just as we have done with higher-order Sobolev spaces, that the experience and maturity gained with first-order problems may be sufficient to clearly see the validity of the following results which is a way to summarize the main points about one problem like our model problem above. It is by no means the most such general result. There might be several variants concerning constraints to be respected either as boundary conditions around $\partial\Omega$, or otherwise. Though, it is a long statement, we believe it is a good way to sum up the fundamental facts.

We are given a density like F in (8.9) which is measurable in the variable $\boldsymbol{x}$, and continuous in $(u, \mathbf{u}, \mathbf{U})$. We consider the corresponding variational problem

$$\text{Minimize in } u \in \mathcal{A} \subset W^{2,p}(\Omega): \quad I(u) = \int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}), \nabla^2 u(\boldsymbol{x}))\, d\boldsymbol{x},$$

where we have added the restriction set $\mathcal{A}$, typically incorporating boundary conditions for feasible functions.

Theorem 8.5

1. *Suppose that*
   (a) *there is a constant* $c > 0$ *and exponent* $p > 1$ *such that*
   $$c(|\mathbf{U}|^p - 1) \leq F(\boldsymbol{x}, u, \mathbf{u}, \mathbf{U}),$$
   or more in general
   $$c\left(\|u\|_{W^{2,p}(\Omega)} - 1\right) \leq I(u);$$
   (b) *the function*
   $$\mathbf{U} \to F(\boldsymbol{x}, u, \mathbf{u}, \mathbf{U})$$
   is convex, for every fixed $(\boldsymbol{x}, u, \mathbf{u})$*;*
   (c) *the class* $\mathcal{A}$ *is weakly closed as a subset of* $W^{2,p}(\Omega)$*.*

   Then our variational problem admits optimal solutions.
2. *Suppose the integrand*
   $$F(u, \mathbf{u}, \mathbf{U}) : \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^{N\times N} \to \mathbb{R}$$
   is continuously differentiable with partial derivatives
   $$F_u(u, \mathbf{u}, \mathbf{U}), \quad F_{\mathbf{u}}(u, \mathbf{u}, \mathbf{U}), \quad F_{\mathbf{U}}(u, \mathbf{u}, \mathbf{U})$$
   and, for some constant $C > 0$*,*
   $$|F(u, \mathbf{u}, \mathbf{U})| \leq C(|\mathbf{U}|^2 + |\mathbf{u}|^2 + |u|^2), \quad |F_u(u, \mathbf{u}, \mathbf{U})| \leq C(|\mathbf{U}| + |\mathbf{u}| + |u|),$$
   $$|F_{\mathbf{u}}(u, \mathbf{u}, \mathbf{U})| \leq C(|\mathbf{U}| + |\mathbf{u}| + |u|), \quad |F_{\mathbf{U}}(u, \mathbf{u}, \mathbf{U})| \leq C(|\mathbf{U}| + |\mathbf{u}| + |u|).$$
   If a function $v \in H^2_0(\Omega)$*, with* $v - u_0 \in H^2_0(\Omega)$*, is a minimizer for the problem*
   $$\textit{Minimize in } u \in H^2(\Omega): \quad I(u) = \int_\Omega F(u(\boldsymbol{x}), \nabla u(\boldsymbol{x}), \nabla^2 u(\boldsymbol{x}))\, d\boldsymbol{x}$$
   under
   $$u - u_0 \in H^2_0(\Omega), \quad u_0 \in H^2(\Omega), \textit{ given},$$

then

$$\int_\Omega \Big[F_u(v(\boldsymbol{x}), \nabla v(\boldsymbol{x}), \nabla^2 v(\boldsymbol{x}))V(\boldsymbol{x}) + F_{\mathbf{u}}(v(\boldsymbol{x}), \nabla v(\boldsymbol{x}), \nabla^2 v(\boldsymbol{x})) \cdot \nabla V(\boldsymbol{x})$$
$$+ F_{\mathbf{U}}(v(\boldsymbol{x}), \nabla v(\boldsymbol{x}), \nabla^2 v(\boldsymbol{x})) : \nabla^2 V(\boldsymbol{x}) \Big] \, d\boldsymbol{x} = 0$$

for every $V \in H_0^1(\Omega)$, i.e. v is a weak solution of the fourth-order problem

$$\text{div}[\text{div}\, F_{\mathbf{U}}(v, \nabla v, \nabla^2 v)] - \text{div}\, F_{\mathbf{u}}(v, \nabla v, \nabla^2 v) + F_u(v, \nabla v, \nabla^2 v) = 0 \text{ in } \Omega,$$
$$v = u_0, \nabla v = \nabla u_0 \text{ on } \partial\Omega.$$

3. *Assume the integrand $F(x, u, \mathbf{u}, \mathbf{U})$ and its partial derivatives*

$$F_u(u, \mathbf{u}, \mathbf{U}), \quad F_{\mathbf{u}}(u, \mathbf{u}, \mathbf{U}), \quad F_{\mathbf{U}}(u, \mathbf{u}, \mathbf{U})$$

enjoy the properties of the previous item, and, in addition, it is convex with respect to triplets $(u, \mathbf{u}, \mathbf{U})$. If a certain function $v \in H^2(\Omega)$, complying with $v - u_0 \in H_0^2(\Omega)$ and u_0, prescribed, is a weak solution of the corresponding fourth-order Euler-Lagrange equation just indicated in the previous item, then v is a minimizer of the associated variational problem.

4. *If in addition to all of the hypotheses indicated in each particular situation, the integrand F is strictly convex in $(u, \mathbf{u}, \mathbf{U})$, there is a unique minimizer and a unique weak solution to the Euler-Lagrange equation, and they are the same function.*

We will be contented with explicitly looking at the most important such example: the bi-laplacian.

Example 8.4 For a regular, bounded domain $\Omega \subset \mathbb{R}^N$, and a given function $u_0 \in H^2(\Omega)$, we want to briefly study the second-order variational problem

$$\text{Minimize in } u \in H^2(\Omega): \quad I(u) = \frac{1}{2}\int_\Omega |\Delta u(\boldsymbol{x})|^2 \, d\boldsymbol{x}$$

under the typical Dirichlet boundary conditions

$$u - u_0 \in H_0^2(\Omega).$$

Recall that

$$\Delta u(\boldsymbol{x}) = \sum_{i=1}^N \frac{\partial^2 u}{\partial x_i^2}(\boldsymbol{x}),$$

and, hence, the integrand we are dealing with is

$$F(\boldsymbol{x}, u, \mathbf{u}, \mathbf{U}) = \frac{1}{2}(\operatorname{tr}\mathbf{U})^2.$$

It is elementary to realize that the associated Euler-Lagrange problem is concerned with the bi-laplacian or bi-harmonic operator

$$\Delta^2 u = 0 \text{ in } \Omega, \quad u = u_0, \frac{\partial u}{\partial \mathbf{n}} = \frac{\partial u_0}{\partial \mathbf{n}} \text{ on } \partial\Omega. \tag{8.10}$$

By introducing $v = u - u_0 \in H_0^2(\Omega)$, so that $u = v + u_0$, and resetting notation $v \mapsto u$, we see that we can work instead, without loss of generality, with a variational problem with integrand

$$F(\boldsymbol{x}, u, \mathbf{u}, \mathbf{U}) = \frac{1}{2}(\operatorname{tr}\mathbf{U})^2 + a(\boldsymbol{x}) \operatorname{tr}\mathbf{U} + F_0(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x})) \tag{8.11}$$

but vanishing boundary conditions $u \in H_0^2(\Omega)$. For the sake of simplicity, we will simply take $F_0 \equiv 0$, and will take $a(\boldsymbol{x}) \in L^2(\Omega)$. We will, therefore, focus on the optimization problem

$$\text{Minimize in } u(\boldsymbol{x}) \in H_0^2(\Omega): \quad I(u) = \int_\Omega \left(\frac{1}{2}|\Delta u(\boldsymbol{x})|^2 + a(\boldsymbol{x})\Delta u(\boldsymbol{x})\right) d\boldsymbol{x}. \tag{8.12}$$

We will take advantage of this special example to stress the following general fact.

If the integrand F of an integral functional I is coercive and (strictly) convex, then I inherit these properties from F. But an integral functional I may be coercive and strictly convex in its domain of definition while none of these properties be correct for the integrand of I pointwise.

It is clear that the integrand F in (8.11) is not coercive on the variable $\mathbf{U}$ nor strictly convex because some of the entries of the hessian of u do not occur in F. Yet we check below that the functional itself I is both coercive and strictly convex in $H_0^2(\Omega)$.

Proposition 8.3 *The functional $I : H_0^2(\Omega) \to \mathbb{R}$ in (8.12) is coercive, smooth and strictly convex. Consequently, there is a unique minimizer for (8.12) and for its corresponding Euler-Lagrange problem (i.e. for (8.10)).*

Proof The proof only require one important, but elementary fact: for every $u \in H_0^2(\Omega)$, it is true that

$$\int_\Omega |\nabla^2 u(\boldsymbol{x})|^2 \, d\boldsymbol{x} = \int_\Omega \Delta u(\boldsymbol{x})^2 \, d\boldsymbol{x}. \tag{8.13}$$

A more general form of this equality is proved in the next chapter. This identity implies that we can also represent our functional in (8.12) by

$$I(u) = \int_\Omega \left(\frac{1}{2} |\nabla^2 u(\boldsymbol{x})|^2 + a(\boldsymbol{x}) \Delta u(\boldsymbol{x}) \right) d\boldsymbol{x}.$$

But the integrand for this representation is

$$\tilde{F}(x, u, \mathbf{u}, \mathbf{U}) = \frac{1}{2} |\mathbf{U}|^2 + a(\boldsymbol{x}) \operatorname{tr} \mathbf{U},$$

which is point-wise coercive and strictly convex in the variable $\mathbf{U}$. We would conclude by Theorem 8.5.

To show (8.13), notice that through a standard density argument, given that $C_c^\infty(\Omega)$ is, by definition (Definition 7.6), dense in $H_0^2(\Omega)$, it suffices to check that for every pair i, j, and for every $u \in C_c^\infty(\Omega)$, it is true that

$$\int_\Omega \frac{\partial}{\partial x_i} \left(\frac{\partial u}{\partial x_j} \right) \frac{\partial}{\partial x_j} \left(\frac{\partial u}{\partial x_i} \right) d\boldsymbol{x} = \int_\Omega \frac{\partial^2 u}{\partial x_j^2} \frac{\partial^2 u}{\partial x_i^2} d\boldsymbol{x}.$$

This identity is checked immediately after two integration by parts (or just by using the product rule for smooth functions) bearing in mind that boundary values for u and ∇u vanish around $\partial\Omega$. □

We can, of course, consider variational problems of order higher than two.

8.9 Non-existence and Relaxation

We have stressed the fundamental concept of convexity throughout this text. Together with coercivity, they are the two building properties upon which existence of optimal solutions to variational problems is proved. Two natural questions arise: what may happen in problems when one of these two ingredients fail? What behavior is expected when the integrand is not convex; or when the integrand is not coercive? There are two terms that refer precisely to these two anomalies: oscillation, when convexity fail; and concentration, when coercivity does. The first indicates a persistent oscillatory phenomenon that is unavoidable in a minimization process, similar to weak convergence when it is not strong; in the second, the loss of mass at infinity is part of the minimization process, and is associated with lack of coercivity. In some cases, both phenomena may produce a combined effect. We are especially interested in having a first contact with these two phenomena.

Example 8.5 Consider the situation

$$\text{Minimize in } u \in H^1(\mathbf{B}) : \quad \int_{\mathbf{B}} \left[\arctan^2(|\nabla u(\boldsymbol{x})|) + u(\boldsymbol{x})^2\right] d\boldsymbol{x},$$

under $u = 1$ on $\partial\mathbf{B}$. $\mathbf{B}$ is the unit ball in $\mathbb{R}^N$. Though the form of this problem may look a bit artificial, it is designed to convey the concentration effect. We claim that the infimum of the problem vanishes. If this is indeed so, then it is clear that there cannot be a function u realizing it, because of the incompatibility of the functional with the boundary datum. Minimizing sequence then should jump abruptly from zero to meet the boundary datum in small boundary layers around $\partial\mathbf{B}$. Indeed, the following is a minimizing sequence of the problem

$$u_j(\boldsymbol{x}) = f_j(|\boldsymbol{x}|), \quad f_j(x) = \begin{cases} 0, & x \in [0, 1 - 1/j], \\ j(x-1)+1, & x \in [1 - 1/j, 1]. \end{cases}$$

If definitely shows the concentration phenomenon we are talking about.

We plan to spend more time with the oscillatory effect that is a consequence of lack of convexity. To this aim we make the following definition.

Definition 8.2 Let $\phi(\mathbf{u}) : \mathbb{R}^N \to \mathbb{R}$ be continuous. We define its convex hull as the function

$$\phi^{\#}(\mathbf{u}) = \inf\left\{\frac{1}{|\mathbf{B}|}\int_{\mathbf{B}} \phi(\nabla u(\boldsymbol{x}))\, d\boldsymbol{x} : u \in C^\infty(\mathbf{B}), u(\boldsymbol{x}) = \mathbf{u}\cdot\boldsymbol{x} \text{ on } \partial\mathbf{B}\right\}.$$

$\mathbf{B}$ is the unit ball in $\mathbb{R}^N$.

A first immediate information is the following.

Proposition 8.4

1. *We always have*

$$\mathbf{C}\phi \le \phi^{\#} \le \phi.$$

2. *If ϕ is convex, then $\phi = \phi^{\#}$.*

Proof The inequality $\phi^\sharp \le \phi$ is trivial because the linear function $u(\boldsymbol{x}) = \mathbf{u}\cdot\mathbf{x}$ is, of course, a valid function to compete for $\phi^\sharp(\mathbf{u})$.

Let u be one of the competing functions in the definition of $\phi^\sharp$ so that $u(\boldsymbol{x}) = \mathbf{u}\cdot\boldsymbol{x}$ on $\partial\mathbf{B}$. Define a measure through the formula

$$\langle\psi, \mu_u\rangle = \frac{1}{|\Omega|}\int_\Omega \psi(\nabla u(\boldsymbol{x}))\, d\boldsymbol{x}.$$

It is clear that every such μ_u is a probability measure, compactly supported in $\mathbb{R}^N$ and barycenter $\mathbf{u}$. By Jensen's inequality, since $\mathbf{C}\phi$ is convex and μ_u, a probability measure with first moment $\mathbf{u}$,

$$\mathbf{C}\phi(\mathbf{u}) \leq \frac{1}{|\mathbf{B}|}\int_{\mathbf{B}} \mathbf{C}\phi(\nabla u(\boldsymbol{x}))\,d\boldsymbol{x} \leq \frac{1}{|\mathbf{B}|}\int_{\mathbf{B}} \phi(\nabla u(\boldsymbol{x}))\,d\boldsymbol{x}.$$

By taking the infimum in u, we conclude that $\mathbf{C}\phi \leq \phi^\sharp$. □

The fundamental and surprising property is the following.

Theorem 8.6 *The function $\phi^\#$ is always convex. Consequently,* $\mathbf{C}\phi \equiv \phi^\#$.

Proof As in the proof of the last proposition, we put

$$\mathbf{F} = \{\mu_u : u \in C^\infty(\mathbf{B}), u(\boldsymbol{x}) = \mathbf{u}\cdot\boldsymbol{x} \text{ on } \partial\mathbf{B}\},$$

and $\mathbf{G}$, the set of all probability measures with first moment $\mathbf{u}$ and compact support in $\mathbb{R}^N$. It is therefore clear that $\mathbf{F} \subset \mathbf{G}$. We will use Corollary 3.6 to show that in fact $\mathbf{G} \subset \overline{\text{co}(\mathbf{F})}$ in the Banach space $C^0(\mathbf{B})$ and its dual, which was identified in Example 7.3 of Chap. 2. Though we did not prove this fact, it is however clear that both sets $\mathbf{F}$ and $\mathbf{G}$, being subsets of probability measures, are subsets of the dual of $C^0(\mathbf{B})$. Let then suppose that for a continuous function ψ, a probability measure $\mu \in \mathbf{G}$, and a real number ρ, we have that

$$\langle\psi, \mu\rangle + \rho < 0.$$

This implies that ψ must be less than $-\rho$ somewhere in $\mathbf{B}$, for otherwise the previous inequality would be impossible. By continuity, this same inequality $\psi < -\rho$ still holds in a certain ball $\mathbf{B}_r(\mathbf{u})$. In this case we can find, for instance, a continuous radial function

$$u(\boldsymbol{x}) = u(|\boldsymbol{x}|), \quad \nabla u(\boldsymbol{x}) = u'(|\boldsymbol{x}|)\frac{\boldsymbol{x}}{|\boldsymbol{x}|},$$

with

$$u(t) : [0, 1] \to \mathbb{R}, \quad u(1) = 0, u' \leq r,$$

in such a way that

$$\mathbf{u}\cdot\boldsymbol{x} + u(\boldsymbol{x}) \in \mathbf{F}, \quad \text{supp}(\nabla u) \subset \mathbf{B}_r(\mathbf{u}).$$

It is then elementary to realize that for such u, the corresponding μ_u is an element of $\mathbf{F}$ and, by construction,

$$\langle \psi, \mu_u \rangle + \rho < 0.$$

□

A fundamental interpretation of the last theorem leads to the remarkable conclusion that the two numbers

$$\inf \left\{ \int_{\mathbf{B}} \mathbf{C}\phi(\nabla u(\boldsymbol{x}))\, d\boldsymbol{x} : u \in C^\infty(\mathbf{B}), u(\boldsymbol{x}) = \mathbf{u} \cdot \boldsymbol{x} \text{ on } \partial\mathbf{B} \right\}$$

and

$$\inf \left\{ \int_{\mathbf{B}} \phi(\nabla u(\boldsymbol{x}))\, d\boldsymbol{x} : u \in C^\infty(\mathbf{B}), u(\boldsymbol{x}) = \mathbf{u} \cdot \boldsymbol{x} \text{ on } \partial\mathbf{B} \right\},$$

are equal because both numbers are equal to $|\mathbf{B}|\mathbf{C}\phi(\mathbf{u})$. This is the simplest example of what in the jargon of the Calculus of Variations is called a relaxation theorem. Since this is a more advanced topic, we just state here a more general such fact for the scalar case, and leave a more complete discussion for a more specialized treatise.

Let

$$F(\boldsymbol{x}, u, \mathbf{u}) : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$$

be measurable with respect to $\boldsymbol{x}$, and continuous with respect to pairs $(u, \mathbf{u})$ and such that there is $p > 1$ and $C > c > 0$ with

$$c(|\mathbf{u}|^p - 1) \le F(\boldsymbol{x}, u, \mathbf{u}) \le C(|\mathbf{u}|^p + 1)$$

for every triplet $(\boldsymbol{x}, u, \mathbf{u})$. As usual, we put

$$I(u) = \int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}.$$

Let $\mathbf{C}F(\boldsymbol{x}, u, \mathbf{u})$ stand for the convexification of $F(\boldsymbol{x}, u, \cdot)$ for each fixed pair $(\boldsymbol{x}, u)$. Then Theorem 8.2 can be applied to the (convexified) functional

$$\mathbf{C}I(u) = \int_\Omega \mathbf{C}F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x},$$

and for given $u_0 \in W^{1,p}(\Omega)$ there is, at least, one minimizer v of the problem.

Theorem 8.7 *Under the indicated hypotheses, we have*

$$\mathbf{C}I(v) = \min\{\mathbf{C}I(u) : u - u_0 \in W^{1,p}(\Omega)\} = \inf\{I(u) : u - u_0 \in W^{1,p}(\Omega)\}.$$

What this statement is claiming is that the problem for the non-convex integrand F may not admit a minimizer for lack of convexity, but if we replace it by its convexification, we recover minimizers and the value of the infimum has not changed.

There are several other important ways to show the convexity of $\phi^\sharp$. These are important for the proof of the previous statement even in a vector situation. As such, they need to be known by anyone interested in deepening his insight into the Calculus of Variation. Since this is a primer on variational techniques, we do not treat them, but just mention them.

The first one is based on the classical Carathéodory theorem, and a funny construction of a Lipschitz function whose graph is some sort of pyramid. Recall the concept of the convexification of a set in Definition 3.4.

Theorem 8.8 *For every set* $\mathbf{C} \subset \mathbb{R}^N$*, we always have*

$$\mathrm{co}(\mathbf{C}) = \left\{ \sum_{i=0}^{N} t_i \mathbf{u}_i : t_i \geq 0, \sum_{i=0}^{N} t_i = 1, \mathbf{u}_i \in \mathbf{C} \right\}.$$

The important point here is that in spite of limiting, to $N+1$, the number of terms in convex combinations with elements of $\mathbf{C}$, yet one does not loose any vector in $\mathrm{co}(\mathbf{C})$.[1]

Proposition 8.5 *Let*

$$\mathbf{u} = \sum_{i=0}^{N} t_i \mathbf{u}_i : t_i \geq 0, \sum_{i=0}^{N} t_i = 1, \mathbf{u}_i \in \mathbb{R}^N.$$

Then there is a simplex $\mathbb{X} \subset \mathbb{R}^N$ *and a Lipschitz function* $u(\boldsymbol{x}) : \mathbb{X} \to \mathbb{R}$ *such that*

$$u(\boldsymbol{x}) = \mathbf{u} \cdot \boldsymbol{x} \text{ on } \partial\mathbb{X},$$

and

$$|\{\nabla u \notin \{\mathbf{u}_i : i = 0, 1, \ldots, N\}\}| = 0.$$

By putting together these two results, one can show, much more explicitly than we have done above, Theorem 8.6.

The second one involves a fundamental construction whose interest goes well beyond our scalar problems here.

[1] We are using the same letter $\mathbf{C}$ for two different things: a set in $\mathbb{R}^N$ and to indicate the convexification of a function. We hope not to lead to any confusion as the context makes clear, we believe, we are referring to in each case

Lemma 8.2 *Let Ω be a bounded, regular domain, and $\mathbf{u}_i \in \mathbb{R}^N$, $i = 1, 0$, $t \in (0, 1)$. Put $\mathbf{u} = t\mathbf{u}_1 + (1 - t)\mathbf{u}_0$. There is a sequence of functions $u_j(\boldsymbol{x}) : \Omega \to \mathbb{R}$ with the following properties:*

1. *$u_j(\boldsymbol{x}) = \mathbf{u} \cdot \boldsymbol{x}$ on $\partial\Omega$;*
2. *$\{u_j\}$ is bounded sequence (of Lipschitz functions) in $W^{1,\infty}(\Omega)$;*
3. *as $j \to \infty$,*

$$|\{\boldsymbol{x} \in \Omega : \nabla u_j(\boldsymbol{x}) \notin \{\mathbf{u}_1, \mathbf{u}_0\}\}| \to 0.$$

8.10 Exercises

1. Consider the variational problem

$$\text{Minimize in } u \in W^{1,p}(\Omega) : \quad \int_\Omega F(u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}$$

where the integrand F is assumed to comply with hypotheses in Theorem 8.2, and no condition is assumed on $\partial\Omega$. Argue that the problem admits an optimal solution, and derive the so-called natural boundary condition for the problem.

2. Show that for every

$$\mathbf{F}(\boldsymbol{x}) : \Omega \subset \mathbb{R}^N \to \mathbb{R}^N, \quad \mathbf{F} \in L^2(\Omega; \mathbb{R}^N),$$

$f(\boldsymbol{x}) \in L^2(\Omega)$, and $u_0 \in H^1(\Omega)$, there is a unique solution $u \in H^1(\Omega)$ solution of the variational problem

$$\text{Minimize in } v \in H^1(\Omega) : \quad \int_\Omega \left[\frac{1}{2}|\nabla v(\boldsymbol{x}) - \mathbf{F}(\boldsymbol{x})|^2 + f(\boldsymbol{x})v(\boldsymbol{x})\right] d\boldsymbol{x}$$

under the boundary condition $v - u_0 \in H_0^1(\Omega)$. Write with care the associated Euler-Lagrange equation.

3. Study the problem

$$\text{Minimize in } u \in H^1(\Omega) :$$
$$\int_\Omega \left[\frac{1}{2}\nabla u(\boldsymbol{x})^T \mathbf{A}(\boldsymbol{x})\nabla u(\boldsymbol{x}) + \frac{1}{2}u(\boldsymbol{x})^2 + f(\boldsymbol{x})u(\boldsymbol{x})\right] d\boldsymbol{x}$$

under no boundary condition whatsoever where the matrix field $\mathbf{A}$ enjoys the usual conditions to guarantee convexity and coercivity. By looking at conditions of optimality, derive the corresponding natural boundary conditions that minimizers of this problem should comply with. Based on this analysis, try

to write down a variational problem that formally would yield a (weak) solution of the problem

$$\Delta u = 0 \text{ in } \Omega, \quad \frac{\partial u}{\partial \boldsymbol{n}} = h \text{ on } \partial\Omega,$$

where $\boldsymbol{n}$ is the unit, outer normal to $\partial\Omega$.

4. Consider the variational problem associated with the functional

$$\int_\Omega \left(\frac{a}{2}|\nabla u|^2 + \nabla u \cdot \mathbf{F}\, \nabla u \cdot \mathbf{G}\right) d\boldsymbol{x}$$

under usual Dirichlet boundary conditions. Give conditions on the two fields

$$\mathbf{F}(\boldsymbol{x}), \mathbf{G}(\boldsymbol{x}) : \Omega \subset \mathbb{R}^N \to \mathbb{R}^N,$$

so that the problem admits a unique solution. Explore the form of the Euler-Lagrange equation.

5. Argue if the non-linear PDE

$$-\operatorname{div}(a\nabla u + b|\nabla u|\mathbf{F}) = 0$$

can be the Euler-Lagrange equation of a certain functional for suitable constant (or functions) a and b, and field $\mathbf{F}$. Find it if the answer is affirmative.

6. Consider the family of radial variational problems with integrands of the form

$$F(\mathbf{u}) = f(|\mathbf{u}|)$$

for $f : [0, \infty) \to \mathbb{R}$. Explore further conditions on f guaranteeing existence of minimizers under standard Dirichlet boundary conditions.

7. Obstacle problems. Show that the problem

$$\text{Minimize in } u(\boldsymbol{x}) \in H_0^1(\Omega) : \quad \frac{1}{2}\int_\Omega |\nabla u(\boldsymbol{x})|^2\, d\boldsymbol{x}$$

under

$$u(\boldsymbol{x}) \geq \phi(\boldsymbol{x}) \text{ in } \Omega$$

admits a unique minimizer, provided ϕ, the obstacle, is a continuous function that is strictly negative around $\partial\Omega$. The set where $u = \phi$ is known as the coincidence set , and its unknown boundary as the corresponding free boundary.

8. Consider a variational problem of the form

$$\int_{\Omega}\left[\frac{1}{2}|\nabla u(\boldsymbol{x})|^2 + \nabla u(\boldsymbol{x})\cdot \mathbf{F}(\boldsymbol{x})\right] d\boldsymbol{x}$$

under a typical Dirichlet condition around $\partial\Omega$, where the field $\mathbf{F}$ is divergence-free $\operatorname{div}\mathbf{F} = 0$ in Ω. Explore the corresponding Euler-Lagrange equation, and try to understand such behavior.

9. Let $\mathbf{A}(\boldsymbol{x})$ be a matrix-field in Ω, not necessarily symmetric. Argue that the two quadratic variational problems with identical linear and zero-th order parts but with quadratic parts

$$\mathbf{u}^T\mathbf{A}(\boldsymbol{x})\mathbf{u}, \quad \mathbf{u}^T\frac{1}{2}(\mathbf{A}(\boldsymbol{x}) + \mathbf{A}(\boldsymbol{x})^T)\mathbf{u}$$

are exactly the same. Conclude that the Euler-Lagrange equation of such a quadratic problem always yield a symmetric problem.

10. Let $\Omega \subset \mathbb{R}^N$ be a bounded, regular domain with a unit, outer normal field $\boldsymbol{n}$ on $\partial\Omega$, and consider the space $\mathbb{H} \equiv L^2_{div}(\Omega)$ of Exercise 7 of the last chapter.

    (a) Use the Lax-Milgram theorem for the bilinear form

    $$\int_{\Omega}\mathbf{F}(\boldsymbol{x})^T\mathbf{A}(\boldsymbol{x})\mathbf{F}(\boldsymbol{x})\, d\boldsymbol{x}, \quad \mathbf{F} \in \mathbb{H},$$

    for a coercive, continuous matrix field $\mathbf{A}(\boldsymbol{x}) : \Omega \to \mathbb{R}^{N\times N}$.

    (b) Derive optimality conditions for the variational problem

    $$\text{Minimize in } \mathbf{F} \in \mathbb{H} : \int_{\Omega}\left[\frac{1}{2}\mathbf{F}(\boldsymbol{x})^T\mathbf{A}(\boldsymbol{x})\mathbf{F}(\boldsymbol{x}) + \mathbf{F}(\boldsymbol{x})\cdot\mathbf{G}(\boldsymbol{x})\right] d\boldsymbol{x},$$

    under

    $$\mathbf{F}(\boldsymbol{x})\cdot\boldsymbol{n}(\boldsymbol{x}) = f(\boldsymbol{x}) \text{ on } \partial\Omega.$$

11. Consider the variational problem for the functional

$$\int_{\Omega}\frac{1}{2}|u_{x_1}(\boldsymbol{x}) + a(\boldsymbol{x})u_{x_2}(\boldsymbol{x})|^2\, d\boldsymbol{x}$$

where

$$a \in L^2(\Omega), \quad \Omega \subset \mathbb{R}^2, \quad \boldsymbol{x} = (x_1, x_2).$$

(a) Explore if one can apply directly the fundamental existence theorem of this chapter under standard Dirichlet conditions on $\partial\Omega$.

(b) Perturb the previous functional in the form

$$\int_\Omega \left[\frac{1}{2}|u_{x_1}(\boldsymbol{x}) + a(\boldsymbol{x})u_{x_2}(\boldsymbol{x})|^2 + \frac{\epsilon}{2}u_{x_2}(\boldsymbol{x})^2\right] d\boldsymbol{x}$$

for a small positive, parameter ϵ. Show that this time there is a unique minimizer u_ϵ of the problem.

(c) What are the changes if we start instead with a functional

$$\int_\Omega \frac{1}{2}|u_{x_1}(\boldsymbol{x}) + a(u(\boldsymbol{x}))u_{x_2}(\boldsymbol{x})|^2\, d\boldsymbol{x}$$

where this time a is a real function as regular as necessary?

12. As in Exercise 1, consider the variational problem

$$\text{Minimize in } u \in W^{1,p}(\Omega): \quad \int_\Omega F(u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}$$

where the integrand F is assumed to comply with hypotheses in Theorem 8.2, and we demand $u = u_0$ only on a given portion $\Gamma \subset \partial\Omega$ of the boundary. Argue that the problem admits an optimal solution, and derive the corresponding optimality conditions.

13. Argue what are the optimal solutions of the variational problem

$$\text{Minimize in } u \in L^\infty(\Omega): \quad \int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}))\, d\boldsymbol{x}$$

under

$$\int_\Omega u(\boldsymbol{x})\, d\boldsymbol{x} = |\Omega| u_0,$$

where u_0 belongs to the range of u.

14. Among all surfaces $\mathbb{S}_u$ that are the graph of a function u of two variables

$$u(x, y) : \Omega \to \mathbb{R}, \quad \Omega \subset \mathbb{R}^2,$$

complying with some prescribed boundary values given by a specific function $u_0(x, y)$,

$$u(x, y) = u_0(x, y), \quad (x, y) \in \partial\Omega,$$

write the functional furnishing the flux of a vector field

$$\mathbf{F}(x, y, z) : \mathbb{R}^3 \to \mathbb{R}^3$$

through $\mathbb{S}_u$. Argue if there is an optimal such surface, and examine the corresponding equation of optimality.

15. Consider the functional

$$I(u) = \int_\Omega \frac{1}{2}(\det \nabla^2 u(\boldsymbol{x}))^2 \, d\boldsymbol{x}, \quad u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}.$$

(a) Check that the function

$$F(\mathbf{A}) = \frac{1}{2} \det \mathbf{A}^2, \quad \mathbf{A} \in \mathbb{R}^{2\times 2},$$

is not convex, and therefore, our main result for second-order problems Theorem 8.5 cannot, in principle, be applied.

(b) Calculate, nonetheless, the corresponding Euler-Lagrange equation.

16. Consider the functional

$$\int_\Omega \frac{\alpha_1 \chi(\boldsymbol{x}) + \alpha_0(1 - \chi(\boldsymbol{x}))}{2} |\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x}$$

where χ and $1 - \chi$ are the characteristic function of two smooth subsets Ω_1 and Ω_0 sharing a smooth boundary Γ within Ω.

(a) Show that there is a unique optimal solution for $u - u_0 \in H_0^1(\Omega)$, provided that α_1, α_0 are non-negative.

(b) Use optimality conditions to describe u with the so-called transmission condition through Γ.

17. Consider the quadratic functional

$$\int_\Omega \frac{a_\epsilon(\boldsymbol{x})}{2} |\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x}$$

with

$$a_\epsilon \rightharpoonup a_0 \text{ in } \Omega, \quad 0 < c \le a_\epsilon, a_0 \le C.$$

(a) Show that there is a unique solution u_ϵ under a typical Dirichlet boundary condition around $\partial\Omega$.

(b) Show that there is a function $u_0 \in H^1(\Omega)$ such that $u_\epsilon \rightharpoonup u_0$ in $H^1(\Omega)$ as $\epsilon \to 0$.

(c) Suppose that for each $v \in H_0^1(\Omega)$ one could find a sequence $\{v_\epsilon\} \subset H_0^1(\Omega)$ such that

$$a_\epsilon \nabla v_\epsilon \to a_0 \nabla v \text{ in } L^2(\Omega).$$

Show that the limit function u_0 is the minimizer of the limit quadratic functional

$$\int_\Omega \frac{a_0(\boldsymbol{x})}{2} |\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x}.$$

In this context, we say the initial quadratic functional Γ-converges to this last quadratic functional.

18. Let Ω be an open subset of $\mathbb{R}^N$ and $u \in W^{1,p}(\Omega)$. Consider the variational problem

$$\text{Minimize in } U \in W^{1,p}(\mathbb{R}^N) : \quad \|U\|_{W^{1,p}(\mathbb{R}^N \setminus \Omega)} + \|U - u\|_{W^{1,p}(\Omega)}.$$

(a) Show that there is a unique minimizer $\hat{u} \in W^{1,p}(\mathbb{R}^N)$.
(b) Suppose that Ω is so regular that the divergence theorem holds. Deduce the corresponding transmission condition as in Exercise 16 above.
(c) Explore the effect of using a positive parameter ϵ to change the functional to

$$\|U\|_{W^{1,p}(\mathbb{R}^N \setminus \Omega)} + \frac{1}{\epsilon} \|U - u\|_{W^{1,p}(\Omega)}.$$

Suppose that u admits an extension to U in the sense that

$$U \in W^{1,p}(\mathbb{R}^N), \quad U = u \text{ in } \Omega.$$

If $\hat{u}_\epsilon$ is the unique minimizer for each ϵ, what behavior would you expect for $\hat{u}_\epsilon$ as $\epsilon \to 0$?

19. This exercise is similar to the previous one for the particular case $p = 2$. Let Ω be an open subset of $\mathbb{R}^N$ and $u \in H^1(\Omega)$. Consider the variational problem

$$I_\epsilon(v) = \frac{1}{2} \int_{\mathbb{R}^N} \left[\chi_{\mathbb{R}^N \setminus \Omega}(\boldsymbol{x}) |\nabla v(\boldsymbol{x})|^2 + \frac{1}{\epsilon} \chi_\Omega(\boldsymbol{x}) |\nabla v(\boldsymbol{x}) - \nabla u(\boldsymbol{x})|^2 \right] d\boldsymbol{x},$$

over the full space $v \in H^1(\mathbb{R}^N)$.

(a) Prove that there is a unique minimizer $v_\epsilon \in H^1(\mathbb{R}^N)$.

(b) Write down the weak formulation of the underlying Euler-Lagrange equation of optimality, and interpret it as a linear operation

$$\mathcal{E}_\epsilon : H^1(\Omega) \to H^1(\mathbb{R}^N), \quad u \mapsto v_\epsilon = \mathcal{E}_\epsilon u.$$

(c) Let m_ϵ be the value of the minimum, and suppose that it is a uniformly bounded sequence of numbers. Use the Banach-Steinhauss principle to show that there is limit operator $\mathcal{E} : H^1(\Omega) \to H^1(\mathbb{R}^N)$ which is linear and continuous.

(d) Interpret the operation $\mathcal{E} : H^1(\Omega) \to H^1(\mathbb{R}^N)$ as an extension operator.

20. Let $\mathbf{F}(\boldsymbol{x})$ be a tangent vector field to $\partial\Omega$ that can be assumed as smooth as necessary. Explore how to deal with the variational problem

$$\text{Minimize in } u : \quad \int_\Omega \frac{1}{2}|\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x}$$

subject to

$$\mathbf{F} \cdot \nabla u = 0 \text{ on } \partial\Omega.$$

Can it be done in a consistent way? Is it a well-posed problem?

21. For the following examples, explore if the basic existence theorem, under Dirichlet boundary conditions, can be applied and write the underlying PDE of optimality. Argue if optimality conditions hold.

$$\int_\Omega \frac{1}{2}\left(\left|\frac{\partial u}{\partial x_1}\right| + \left|\frac{\partial u}{\partial x_2}\right|\right)^2 dx_1\, dx_2,$$

$$\int_\Omega \frac{1}{2}\left(2\left(\frac{\partial u}{\partial x_1}\right)^2 + \left(\frac{\partial u}{\partial x_2}\right)^2 - 2\frac{\partial u}{\partial x_1}\frac{\partial u}{\partial x_2}\right) dx_1\, dx_2,$$

$$\int_\Omega \frac{1}{2}\left|\left(\frac{\partial u}{\partial x_1}\right)^2 + \left(\frac{\partial u}{\partial x_2}\right)^2 - 2\frac{\partial u}{\partial x_1}\frac{\partial u}{\partial x_2}\right| dx_1\, dx_2,$$

$$\int_\Omega \frac{1}{2}\left(\left|\frac{\partial u}{\partial x_1}\right|^4 + \left|\frac{\partial u}{\partial x_2}\right|^4\right)^{1/2} dx_1\, dx_2,$$

$$\int_\Omega \frac{1}{2}\left(2\left(\frac{\partial u}{\partial x_1}\right)^4 + \left(\frac{\partial u}{\partial x_2}\right)^4 - 2\left(\frac{\partial u}{\partial x_1}\frac{\partial u}{\partial x_2}\right)^2\right)^{1/2} dx_1\, dx_2,$$

$$\int_\Omega \frac{1}{2}\left[\left(\left(\frac{\partial u}{\partial x_1}\right)^4 + \left(\frac{\partial u}{\partial x_2}\right)^2\right)^{1/2} + \left(\frac{\partial u}{\partial x_2}\right)^2\right] dx_1\, dx_2,$$

$$\int_\Omega \frac{1}{2}\left[\left(\left(\frac{\partial u}{\partial x_1}\right)^4+\left(\frac{\partial u}{\partial x_2}\right)^2\right)^{1/2}+\left(\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^4\right)^{1/2}\right] dx_1\,dx_2,$$

$$\int_\Omega \left[\frac{1}{2}\left(\frac{\partial u}{\partial x_1}\right)^2+\frac{1}{2}\left(\frac{\partial u}{\partial x_2}\right)^2+7\exp\left(-\left(\frac{\partial u}{\partial x_1}-1\right)^4-\left(\frac{\partial u}{\partial x_2}\right)^2\right)\right] dx_1\,dx_2,$$

$$\int_\Omega \left|\frac{\partial u}{\partial x_1}\right|\left|\frac{\partial u}{\partial x_2}\right| dx_1\,dx_2,$$

$$\int_\Omega \left(1+\left(\frac{\partial u}{\partial x_1}\right)^2\right)^{1/2}\left(1+\left(\frac{\partial u}{\partial x_2}\right)^2\right)^{1/2} dx_1\,dx_2,$$

$$\int_\Omega \left(1+\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right)^{1/2} dx_1\,dx_2.$$

22. Other boundary conditions. In the following three situations, investigate the interplay between the functional and the boundary condition. For the sake of simplicity, take in all three cases the usual quadratic functional

$$\frac{1}{2}\int_\Omega |\nabla u(\boldsymbol{x})|^2\,d\boldsymbol{x}.$$

    (a) Mixed boundary condition. The boundary $\partial\Omega$ of the domain Ω is divided in two sets Γ_0 and Γ_n where a Dirichlet condition $u = u_0$ and a Neumman condition $\nabla u \cdot \mathbf{n} = 0$, respectively, are imposed.
    (b) Robin boundary condition. This time we seek an explicit linear dependence between a Dirichlet and a Neumman condition

$$\nabla u \cdot \mathbf{n} = \gamma u \text{ on } \partial\Omega,$$

    where γ is a non-null constant.
    (c) A mixed boundary condition with a free boundary between the Dirichlet and Neumann contition. Take as boundary condition $u \le u_0$ on $\partial\Omega$.

23. If $\Omega \subset \mathbb{R}^3$, write the Euler-Lagrange equation for the functional

$$E(u) = \frac{1}{2}\int_\Omega |\nabla u(\boldsymbol{x}) \wedge \nabla w(\boldsymbol{x}) - \mathbf{F}(\boldsymbol{x})|^2\,d\boldsymbol{x}$$

where the function w and the field $\mathbf{F}$ are given so that the gradient ∇w of w and $\mathbf{F}$ are parallel at every point $\boldsymbol{x} \in \Omega$

$$\nabla w(\boldsymbol{x}) \parallel \mathbf{F}(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega.$$

$\mathbf{u} \wedge \mathbf{v}$ stands for the cross or vector product in $\mathbb{R}^3$.

Chapter 9
Finer Results in Sobolev Spaces and the Calculus of Variations

9.1 Overview

This chapter focuses on some important issues which arise when one keeps working with Sobolev spaces in variational problems and PDEs. We can hardly cover all of the important topics, but have tried to select those which, we believe, are among the ones that might make up a second round on Sobolev spaces and variational problems.

The first situation we will be dealing with is that of variational problems under additional constraints, other than boundary values, in the form of global, integral conditions. In general terms, we are talking about problems of the form

$$\text{Minimize in } u \in \mathcal{A}: \quad \int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}$$

where the set $\mathcal{A}$ of competing functions is a subset of some $W^{1,p}(\Omega)$, which, in addition to boundary conditions around $\partial\Omega$, incorporates some integral constraints given in the form

$$\int_\Omega \mathbf{F}(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x} \le (=) \boldsymbol{f}$$

where

$$\mathbf{F}(\boldsymbol{x}, u, \mathbf{u}) : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^d, \quad \boldsymbol{f} \in \mathbb{R}^d.$$

The various components of the vector-valued map $\mathbf{F}$ are assumed to be smooth. We have already seen one important such example, though constraints are imposed in a

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4_9

point-wise fashion, in Exercise 29 of the last chapter. For such a situation

$$F = \frac{1}{2}|\mathbf{u}|^2, \quad \mathcal{A} = \{u \in H_0^1(\Omega) : u \geq \psi \text{ in } \Omega\},$$

where the continuous function ψ represents the obstacle. Another fundamental problem is that in which we limit the size of feasible u's in some suitable Lebesgue space. The existence of minimizers for such problems is essentially covered by our results in the previous chapter. We will also look at optimality conditions.

A second important point is that of better understanding the quantitative relationship between the norm of a function u in $W^{1,p}(\Omega)$ and its norm in the Lebesgue space $L^q(\Omega)$, for a domain $\Omega \subset \mathbb{R}^N$. This is directly related to variational problems of the form

$$\int_\Omega \left[\frac{1}{p}|\nabla u(\boldsymbol{x})|^p + \frac{1}{p}|u(\boldsymbol{x})|^p - \lambda \frac{1}{q}|u(\boldsymbol{x})|^q \right] d\boldsymbol{x} \tag{9.1}$$

for different exponents p, q, and positive constant λ. Note that if $\lambda \leq 0$, the unique minimizer of the previous functional in $W_0^{1,p}(\Omega)$ is the trivial function. However, the problem becomes meaningful (non-trivial) if $\lambda > 0$. Indeed, the functional in (9.1) is closely related to the constrained variational problem

$$\text{Minimize in } u \in W^{1,p}(\Omega) : \quad \|u\|^p_{W^{1,p}(\Omega)}$$

under the integral constraint

$$\|u\|_{L^q(\Omega)} = C, \quad C > 0, \text{ given.}$$

To better realize what is at stake in such situation, suppose we take Ω all of space $\Omega = \mathbb{R}^N$. For real, positive r, we put

$$u_r(\boldsymbol{x}) = u(r\boldsymbol{x}), \quad u \in W^{1,p}(\mathbb{R}^N).$$

A simple computation leads to

$$\|\nabla u_r\|_{L^p(\Omega)} = r^{1-N/p}\|\nabla u\|_{L^p(\Omega)}, \quad \|u\|_{L^q(\Omega)} = r^{-N/q}\|u\|_{L^q(\Omega)},$$

for each such r, and exponents p and q. This identities show that $u_r \in W^{1,p}(\mathbb{R}^N)$ for all such r; and, moreover, if the norms of ∇u in $L^p(\mathbb{R}^N)$ and u in $L^q(\mathbb{R}^N)$ are to be comparable for every function u in $W^{1,p}(\mathbb{R}^N)$, then we need to necessarily have that

$$\frac{1}{N} = \frac{1}{p} + \frac{1}{q}, \tag{9.2}$$

for otherwise, the limits for $r \to 0$ or $r \to \infty$ would make invalid the comparison. The condition (9.2) yields the relationship between exponent p for the Sobolev space, exponent q for the Lebesgue space, and the dimension N of the domain. The situation for a different (in particular bounded) domain is much more involved and technical. Indeed, the usual way of looking at the issue is by resorting to the case of the whole space $\mathbb{R}^N$ through the use of an appropriate extension operator Υ defined as

$$\Upsilon : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^N), \quad \|\Upsilon u\|_{W^{1,p}(\mathbb{R}^N)} \leq C_\Omega \|u\|_{W^{1,p}(\Omega)}, \tag{9.3}$$

for a constant C_Ω depending exclusively in the domain Ω (and dimension N). If this is the case, and assuming that we establish first an inequality

$$\|v\|_{L^q(\mathbb{R}^N)} \leq C_{N,p,q} \|v\|_{W^{1,p}(\mathbb{R}^N)},$$

for every $v \in W^{1,p}(\mathbb{R}^N)$, we would have

$$\|u\|_{L^q(\Omega)} \leq \|\Upsilon u\|_{L^q(\mathbb{R}^N)} \leq C_{N,p,q} \|\Upsilon u\|_{W^{1,p}(\mathbb{R}^N)} \leq C_{N,p,q} C_\Omega \|u\|_{W^{1,p}(\Omega)},$$

for every $u \in W^{1,p}(\Omega)$. Though there is, apparently, no full characterization of those domains that admit such extension operators with the fundamental property in (9.3), there are various ways to construct them based on additional smoothness of their boundaries. Most of them require a good deal of technicalities. One way relies on the use of partitions of unity, as well as other more intuitive extension techniques like reflection. This process typically requires C^1-smoothness. Other techniques only demand Lipschitzianity of $\partial\Omega$. In fact, there is a particular case in which such extension procedure is especially easy: from Proposition 7.8, we already know that for functions in $W_0^{1,p}(\Omega)$ the extension by zero off Ω makes the function an element of $W^{1,p}(\mathbb{R}^N)$, or of $W^{1,p}(\tilde{\Omega})$ for any bigger domain $\tilde{\Omega}$ for that matter. We will therefore start with the study of the relationship between the norms in $L^q(\Omega)$ and $W^{1,p}(\Omega)$ for functions in $W_0^{1,p}(\Omega)$. We will also briefly describe a different technique that is sufficient to deal with most of the regular domains, and is based on Definition 7.3 and ODEs techniques.

One of the most studied scalar variational problems corresponds to the integrand

$$F(u, \mathbf{u}) = \frac{1}{2}|\mathbf{u}|^2 + f(u)$$

where the continuous function $f(u)$ is placed into varying sets of assumptions. One could even allow for an explicit $\boldsymbol{x}$-dependence, though for simplicity we will stick to the above model problem. To fix ideas, we could think of f as a polynomial of a certain degree. If $f(u)$ is strictly convex and bounded from below by some constant, then our results in the previous chapter cover the situation, and we can conclude that there is a unique critical function (a unique solution of the underlying

Euler-Lagrange equation) which is the unique minimizer of the problem. If we are willing to move one step further, and be dispensed with these assumptions, then more understanding is required concerning the two basic ingredients of the direct method:

1. the convexity requirement in Theorem 8.2 does not demand the convexity of $f(u)$, since it suffices the convexity of $F(u, \mathbf{u})$ with respect to $\mathbf{u}$;
2. the coercivity condition in that same result (for $p = 2$) might not hold true if $f(u)$ is not bounded from below.

Therefore, there are two main issues to be considered for a functional like

$$I(u) = \int_\Omega \left(\frac{1}{2} |\nabla u(\boldsymbol{x})|^2 + f(u(\boldsymbol{x})) \right) d\boldsymbol{x}. \tag{9.4}$$

One is to understand more general coercivity conditions than those in Theorem 8.2 that still may guarantee the existence of minimizers. The second is to explore under what circumstances this last functional may admit critical functions other than minimizers. Note that the Euler-Lagrange equation for such a functional is

$$-\Delta u + f'(u) = 0 \text{ in } \Omega.$$

The first question leads to investigating further integrability properties of Sobolev functions, while the second pushes in the direction of finding methods to show the existence of critical functions solutions of the above PDE not relying on minimization.

Another final fundamental point we would like to consider is that of the further regularity of weak solutions of PDEs, let them be minimizers or just critical functions of a certain functional. We will concentrate on the following basic problem:

$$\Delta u \in L^2(\Omega) \iff u \in H^2(\Omega).$$

This is a quite delicate topic that we plan to address by means of considering higher-order variational problems. Second-order boundary-value problems are also important and appealing. We have already had a chance to explore the important case of the bi-harmonic operator in the last chapter which is associated with the second-order functional

$$\int_\Omega \frac{1}{2} |\Delta u(\boldsymbol{x})|^2 \, d\boldsymbol{x}.$$

Their nature depends in a very fundamental way on the boundary conditions that are enforced in concrete variational problems.

9.2 Variational Problems Under Integral Constraints

Variational problems in which integral constraints, in addition to boundary values, are to be respected as part of feasibility can be interesting and important. We have already mentioned and proposed one such example in the last chapter. As a matter of fact, the main existence result for such problems comes directly from a direct generalization of Theorem 8.2 which is modeled after Theorem 4.3 that we transcribe next.

Suppose that

$$F(\boldsymbol{x}, u, \mathbf{u}) : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R},$$

is measurable in $\boldsymbol{x}$ and continuous in $(u, \mathbf{u})$. Consider the variational problem

$$\text{Minimize in } u(\boldsymbol{x}) \in \mathcal{A} : \quad I(u) = \int_\Omega F(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x}$$

where the feasible set $\mathcal{A}$ is a subset of $W^{1,p}(\Omega)$.

Theorem 9.1 *If the three conditions:*

1. *I is non-trivial, and coercive in the sense*

$$I(u) \to \infty \text{ whenever } \|u\|_{W^{1,p}(\Omega)} \to \infty;$$

2. *$F(\boldsymbol{x}, u, \cdot)$ is convex for every pair $(\boldsymbol{x}, u)$;*
3. *$\mathcal{A}$ is weakly closed in $W^{1,p}(\Omega)$;*

hold, then the above variational problem admits minimizers.

The proof must be clear at this stage right after the direct method. As usual, the coercivity of the functional I is derived from the coercivity of the integrand F just as in Theorem 8.2. Our emphasis is here is in some typical examples of feasible sets $\mathcal{A}$ complying with the weak closeness in the previous statement.

Proposition 9.1 *Let*

$$\mathbf{F}(\boldsymbol{x}, u, \mathbf{u}) : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^d,$$

be a mapping whose components are measurable in $\boldsymbol{x}$ and continuous on $(u, \mathbf{u})$.

1. *If every component of $\mathbf{F}$ is convex in $\mathbf{u}$, then*

$$\mathcal{A} = \{u \in W_0^{1,p}(\Omega) : \int_\Omega \mathbf{F}(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x} \le \boldsymbol{f}\}$$

is weakly closed for every fixed $\boldsymbol{f} \in \mathbb{R}^d$.

2. *If every component of* $\mathbf{F}$ *is linear in* $\mathbf{u}$ *(in particular for those components of* $\mathbf{F}$ *independent of* $\mathbf{u}$*), then*

$$\mathcal{A} = \{u \in W_0^{1,p}(\Omega) : \int_\Omega \mathbf{F}(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x} = \boldsymbol{f}\}$$

is weakly closed for every fixed $\boldsymbol{f} \in \mathbb{R}^d$.

One can obviously mix the equality/inequality conditions for different components of $\mathbf{F}$ asking for the linearity/convexity of the corresponding components. The proof amounts to the realization that, due to the weak lower semicontinuity, of the family of functionals

$$\int_\Omega \mathbf{F}(\boldsymbol{x}, u(\boldsymbol{x}), \nabla u(\boldsymbol{x}))\, d\boldsymbol{x},$$

the definition of the admissible set $\mathcal{A}$ with inequalities turn out to be weakly closed. For the case of linearity, note that both $\mathbf{F}$ and $-\mathbf{F}$ are convex in $\mathbf{u}$.

As indicated above, the particular case for an obstacle problem

$$\mathcal{A} = \{u \in H_0^1(\Omega) : u \geq \psi \text{ in } \Omega\}$$

falls under the action of the preceding proposition, and so does

$$\mathcal{A} = \{u \in W^{1,p}(\Omega) : \|u\|_{L^p(\Omega)} = k\}$$

for a fixed positive constant k. In both cases, there are unique minimizers for the minimization of the $L^2(\Omega)$-norm and the $L^p(\Omega)$-norm of the gradient, respectively (if $p > 1$).

Optimality for such constrained problems involves the use of multipliers. This can be setup in full generality, but since the most studied examples correspond to very specific situations, we will be contented with looking at the following example.

Example 9.1 Let us investigate the problem

$$\text{Minimize in } u \in H_0^1(\Omega) : \quad \frac{1}{2}\int_\Omega |\nabla u(\boldsymbol{x})|^2\, d\boldsymbol{x}$$

subject to the condition

$$\int_\Omega u(\boldsymbol{x})^2\, d\boldsymbol{x} = 1.$$

It is easy to see how the combination of Theorem 9.1 and Proposition 9.1 yields immediately the existence of a unique global minimizer u_0 for this constrained

variational problem. Let

$$m = \frac{1}{2}\int_{\Omega} |\nabla u_0(\boldsymbol{x})|^2 \, d\boldsymbol{x} > 0.$$

If u is an arbitrary, non-trivial function in $H_0^1(\Omega)$, then it is elementary to realize that

$$\overline{u} \equiv \frac{1}{\|u\|_{L^2(\Omega)}} u$$

is feasible for our constrained variational problem, and hence

$$\frac{1}{2}\int_{\Omega} |\nabla \overline{u}(\boldsymbol{x})|^2 \, d\boldsymbol{x} \geq m,$$

that is

$$\int_{\Omega} \left[\frac{1}{2}|\nabla u(\boldsymbol{x})|^2 - mu(\boldsymbol{x})^2\right] d\boldsymbol{x} \geq 0.$$

This inequality exactly means that u_0 is also a global minimizer for the augmented functional

$$\tilde{I}(u) = \int_{\Omega} \left[\frac{1}{2}|\nabla u(\boldsymbol{x})|^2 - mu(\boldsymbol{x})^2\right] d\boldsymbol{x},$$

because $\tilde{I}(u_0) = 0$. The function u_0 is, then, a solution of the corresponding Euler-Lagrange equation

$$-\Delta u = 2mu \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega.$$

We say that the problem

$$-\Delta u = \lambda u \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega \tag{9.5}$$

is the optimality condition for the initial constrained problem for a certain value for the multiplier $\lambda (= 2m)$. Note how Eq. (9.5) can be interpreted by saying that $u = u_0$ is an eigenfunction of the (negative) Laplace operator corresponding to the eigenvalue λ.

The obstacle problem is different because constraints are enforced in a point-wise manner, and so multipliers would not be vectors but functions. We will not treat this kind of situations here.

9.3 Sobolev Inequalities

One fundamental chapter in Sobolev spaces is that of understanding how much better functions u in $W^{1,p}(\Omega)$ are compared to functions in Lebesgue spaces $L^p(\Omega)$ depending on the exponent p and the dimension N where $\Omega \subset \mathbb{R}^N$. The basic tool is again the fundamental theorem of Calculus used as in the proofs of Propositions 7.3, 7.5 and 7.9. Once again we can write

$$u(\boldsymbol{x}', x) = u(\boldsymbol{x}', y) + \int_y^x \frac{\partial u}{\partial x_N}(\boldsymbol{x}', z)\, dz \tag{9.6}$$

for a.e. $\boldsymbol{x}' \in \pi_N \Omega$, and $x, y \in \mathbb{R}$ so that

$$(\boldsymbol{x}', x), (\boldsymbol{x}', y) \in \Omega.$$

If, as usual, we put

$$\Omega = \{(\boldsymbol{x}', x) \in \mathbb{R}^N : \boldsymbol{x}' \in \pi_N \Omega, x \in J_{\boldsymbol{x}'}\},$$

and

$$y(\boldsymbol{x}') = \inf J_{\boldsymbol{x}'}, \quad (\boldsymbol{x}', y(\boldsymbol{x}')) \in \partial\Omega,$$

then (9.6) becomes

$$|u(\boldsymbol{x}', x)| \leq |u(\boldsymbol{x}', y(\boldsymbol{x}'))| + \int_{J_{\boldsymbol{x}'}} \left| \frac{\partial u}{\partial x_N}(\boldsymbol{x}', z) \right| dz.$$

At this initial stage, we are not especially interested in measuring the effect of the term $|u(\boldsymbol{x}', y(\boldsymbol{x}'))|$ at the boundary, which will take us to introducing spaces of functions over boundaries of sets, so that we will assume, to begin with, that $u \in W_0^{1,p}(\Omega)$. If this is so, then

$$u(\boldsymbol{x}', y(\boldsymbol{x}')) = 0,$$

and

$$|u(\boldsymbol{x})| = |u(\boldsymbol{x}', x)| \leq \int_{J_{\boldsymbol{x}'}} \left| \frac{\partial u}{\partial x_N}(\boldsymbol{x}', z) \right| dz. \tag{9.7}$$

We define the function $u_N(\boldsymbol{x}')$ as the right-hand side in this inequality

$$u_N(\boldsymbol{x}') = \int_{J_{\boldsymbol{x}'}} \left| \frac{\partial u}{\partial x_N}(\boldsymbol{x}', z) \right| dz,$$

and realize that we could redo all of the above computation along any of the coordinate axes $\mathbf{e}_i$ to find

$$|u(\boldsymbol{x})| \leq u_i(\boldsymbol{x}_i'), \quad i = 1, 2, \dots, N, \tag{9.8}$$

where

$$u_i(\boldsymbol{x}_i') \equiv \int_{J_{x_i'}} \left| \frac{\partial u}{\partial x_i}(\boldsymbol{x}_i', z) \right| dz.$$

We would like to profit inequalities (9.8) as much as possible. We know that each right-hand side in (9.8) belongs to L^1 of its domain

$$u_i(\boldsymbol{x}_i') : \pi_i\Omega \to \mathbb{R}, \quad \|u_i\|_{L^1(\pi_i\Omega)} = \left\| \frac{\partial u}{\partial x_i} \right\|_{L^1(\Omega)},$$

provided $u \in W^{1,1}(\Omega)$. But (9.8) also implies

$$|u(\boldsymbol{x})|^{N/(N-1)} \leq \prod_i u_i(\boldsymbol{x}_i')^{1/(N-1)}, \tag{9.9}$$

and we have the following abstract technical fact.

Lemma 9.1 *Suppose we have $N(\geq 2)$ functions $f_i \in L^{N-1}(\pi_i\Omega)$ for open sets $\pi_i\Omega \subset \mathbb{R}^{N-1}$, $i = 1, 2, \dots, N$, and put*

$$f(\boldsymbol{x}) = \prod_i f_i(\pi_i \boldsymbol{x}).$$

Then $f \in L^1(\Omega)$, and

$$\|f\|_{L^1(\Omega)} \leq \prod_i \|f_i\|_{L^{N-1}(\pi_i\Omega)}.$$

Proof The first case $N = 2$ does not require any comment. Take $N = 3$, and put, to better see the structure of the situation,

$$f(x_1, x_2, x_3) = f_1(x_2, x_3) f_2(x_1, x_3) f_3(x_1, x_2),$$

with the three factors on the right being functions in L^2 of the corresponding domains. For a.e. pair $(x_2, x_3) \in \pi_1\Omega$, we see that

$$\int_{J_{\pi_1 x}} |f(x_1, x_2, x_3)|\, dx_1 \leq$$

$$|f_1(x_2, x_3)| \left(\int_{J_{\pi_1 x}} |f_2(x_1, x_3)|^2\, dx_1 \right)^{1/2} \left(\int_{J_{\pi_1 x}} |f_3(x_1, x_2)|^2\, dx_1 \right)^{1/2}.$$

Integration on the variables $(x_2, x_3) \in \pi_1\Omega$, together with Hölder's inequality leads to the desired inequality. The general case proceed by induction (exercise). □

If we go back to our situation in (9.9), and apply to it Lemma 9.1, we conclude that $u \in L^{N/(N-1)}(\Omega)$, and, since,

$$\|u\|_{L^{N/(N-1)}(\Omega)}^{N/(N-1)} = \| |u|^{N/(N-1)} \|_{L^1(\Omega)}$$

we arrive at

$$\|u\|_{L^{N/(N-1)}(\Omega)}^{N/(N-1)} \leq \prod_i \left\| \frac{\partial u}{\partial x_i} \right\|_{L^1(\Omega)}^{1/(N-1)},$$

Finally, because

$$\left\| \frac{\partial u}{\partial x_i} \right\|_{L^1(\Omega)} \leq \|\nabla u\|_{L^1(\Omega;\mathbb{R}^N)},$$

we find

$$\|u\|_{L^{N/(N-1)}(\Omega)} \leq \|\nabla u\|_{L^1(\Omega;\mathbb{R}^N)}.$$

Lemma 9.2 *Let $\Omega \subset \mathbb{R}^N$ be a domain, and $u \in W_0^{1,1}(\Omega)$. Then $u \in L^{N/(N-1)}(\Omega)$, and*

$$\|u\|_{L^{N/(N-1)}(\Omega)} \leq \|\nabla u\|_{L^1(\Omega;\mathbb{R}^N)}.$$

This is the kind of result we are looking for. It asserts that functions in $W_0^{1,1}(\Omega)$ enjoy much better integrability properties since they belong to the more restrictive class of functions $L^{N/(N-1)}(\Omega)$.

9.3.1 The Case of Vanishing Boundary Data

Assume that $\Omega \subset \mathbb{R}^N$ is a domain, and $u \in W_0^{1,p}(\Omega)$. We will divide our discussion in three cases depending on the relationship between the exponent p of integrability of derivatives, and the dimension N of the domain, according to the following classification.

1. The subcritical case $1 \le p < N$.
2. The critical case $p = N$.
3. The supercritical case $p > N$.

We will treat the three cases successively.

9.3.1.1 The Subcritical Case

We place ourselves in the situation where the integrability exponent p is strictly smaller than dimension N. We will be using recursively Lemma 9.2 in the following way.

Suppose $u \in W_0^{1,p}(\Omega)$ with $1 < p < N$. Take a exponent $\gamma > 1$, not yet selected. Consider the function

$$U(\boldsymbol{x}) = |u(\boldsymbol{x})|^{\gamma-1}u(\boldsymbol{x}),$$

and suppose that γ is such that $U \in W_0^{1,1}(\Omega)$. It is easy to calculate that

$$\nabla U(\boldsymbol{x}) = \gamma|u(\boldsymbol{x})|^{\gamma-1}\nabla u(\boldsymbol{x}).$$

If we want this gradient to be integrable ($U \in W_0^{1,1}(\Omega)$), since ∇u belongs to $L^p(\Omega)$, we need, according to Hölder's inequality, that $|u|^{\gamma-1} \in L^{p/(p-1)}(\Omega)$. Given that $u \in L^p(\Omega)$, this forces us to take $\gamma = p$. If we do so, then Lemma 9.2, implies that in fact $U = |u|^{\gamma-1}u$ belongs to $L^{N/(N-1)}(\Omega)$, that is, $u \in L^{pN/(N-1)}(\Omega)$. Once we know this, we can play the same game with a different exponent γ, in fact, with

$$(\gamma-1)\frac{p}{p-1} = \frac{pN}{N-1}.$$

If we so choose γ, the same previous computations lead to another U to which we can apply Lemma 9.2. By the new choice of γ, the factor $|u|^{\gamma-1} \in L^{p/(p-1)}(\Omega)$. The conclusion is that $u \in L^{\gamma N/(N-1)}(\Omega)$. Proceeding recursively, we realize that we can move on with this bootstrap argument to reach a exponent γ such that

$$(\gamma-1)\frac{p}{p-1} = \frac{\gamma N}{N-1},$$

and, in such a case, $u \in L^{\gamma N/(N-1)}(\Omega)$. Some elementary arithmetic yields

$$\gamma = \frac{p(N-1)}{N-p}, \quad u \in L^{Np/(N-p)}(\Omega).$$

Theorem 9.2 *If $u \in W_0^{1,p}(\Omega)$, and $1 \le p < N$, then $u \in L^{Np/(N-p)}(\Omega)$, and*

$$\|u\|_{L^{Np/(N-p)}(\Omega)} \le C\|\nabla u\|_{L^p(\Omega)}$$

for a positive constant C depending on p and N (but independent of Ω, exercise below).

The exponent

$$p^* = \frac{pN}{N-p}, \quad 1 \le p < N,$$

is identified as the critical exponent.

For bounded domains, we can state a more complete result.

Theorem 9.3 *If $\Omega \subset \mathbb{R}^N$ is a bounded domain and $1 \le p < N$, then*

$$W_0^{1,p}(\Omega) \subset L^q(\Omega), \quad q \in [1, p^*],$$

and

$$\|u\|_{L^q(\Omega)} \le C\|\nabla u\|_{L^p(\Omega)},$$

for a constant depending on p, q, N, and Ω.

The proof is a combination of the case $q = p^*$ shown in the previous result, and Hölder's inequality when $q < p^*$ (Ω, bounded).

9.3.1.2 The Critical Case

When $p = N$, notice that the critical exponent $p^* = \infty$. As a matter of fact, in this situation the recursive process above corresponding to the proof of Theorem 9.2 never ends (exercise), and we can conclude that $u \in L^q(\Omega)$ for every q, in case Ω is bounded. This is however different to saying that $u \in L^\infty(\Omega)$. In fact, for $N > 1$, there are functions in $W^{1,N}(\mathbb{R}^N)$ which do not belong to $L^\infty(\Omega)$.

Theorem 9.4 *Every function $u \in W_0^{1,N}(\Omega)$ belongs to every $L^q(\Omega)$ for all finite q.*

9.3.1.3 The Supercritical Case

Let us now deal with the supercritical case $p > N$. To better understand and appreciate the relevance that the power of integrability be strictly larger than dimension, let us recall first the case $N = 1$. The conclusion in this case is exactly inequality (2.8) where we showed easily that

$$|u(x) - u(y)| \le \|u'\|_{L^p(J)}|x - y|^{1-1/p}, \quad p > 1,$$

for $x, y \in J \subset \mathbb{R}$. As matter of fact, our approach to higher dimensional Sobolev spaces relies in the fact that one-dimensional sections along coordinate axes are continuous for a.e. such line. Recall Proposition 7.3 and even more so Proposition 7.4. We proceed in several steps.

Step 1. By Proposition 7.8, we can assume that $u \in W^{1,p}(\mathbb{R}^N)$, so that the domain Ω is all of space. We regard, then, u as defined for every vector $\boldsymbol{x} \in \mathbb{R}^N$.

Step 2. Suppose that $\mathbf{Q}$ is any cube with axes parallel to the coordinate axes, and such that $\mathbf{0} \in \mathbf{Q}$. For any point $\boldsymbol{x} \in \mathbf{Q}$, we have, according to Proposition 7.4, that

$$u(\boldsymbol{x}) - u(\mathbf{0}) = \int_0^1 \nabla u(r\boldsymbol{x}) \cdot \boldsymbol{x}\, dr.$$

If we integrate with respect to $\boldsymbol{x} \in \mathbf{Q}$, we arrive at

$$\int_{\mathbf{Q}} u(\boldsymbol{x})\, d\boldsymbol{x} - |\mathbf{Q}|u(\mathbf{0}) = \int_{\mathbf{Q}} \int_0^1 \nabla u(r\boldsymbol{x}) \cdot \boldsymbol{x}\, dr\, d\boldsymbol{x}.$$

If we interchange the order of integration in the last integral, we can also write

$$\int_{\mathbf{Q}} u(\boldsymbol{x})\, d\boldsymbol{x} - |\mathbf{Q}|u(\mathbf{0}) = \int_0^1 \int_{\mathbf{Q}} \nabla u(r\boldsymbol{x}) \cdot \boldsymbol{x}\, d\boldsymbol{x}\, dr.$$

We can perform the change of variables

$$\boldsymbol{y} = r\boldsymbol{x}, \quad d\boldsymbol{y} = r^N d\boldsymbol{x},$$

in the inner integral, for each $r \in (0, 1)$ fixed, and find that

$$\int_{\mathbf{Q}} u(\boldsymbol{x})\, d\boldsymbol{x} - |\mathbf{Q}|u(\mathbf{0}) = \int_0^1 \frac{1}{r^{N+1}} \int_{r\mathbf{Q}} \nabla u(\boldsymbol{y}) \cdot \boldsymbol{y}\, d\boldsymbol{y}\, dr.$$

For the inner integral, after using the classic Cauchy-Schwarz inequality and Hölder's inequality, bearing in mind that $r\mathbf{Q} \subset \mathbf{Q}$ for $r \in (0, 1)$,

$$\int_{r\mathbf{Q}} \nabla u(\boldsymbol{y}) \cdot \boldsymbol{y}\, d\boldsymbol{y} \le \|\nabla u\|_{L^p(\mathbf{Q};\mathbb{R}^N)} r^{1+N(p-1)/p} s^{1+N(p-1)/p}$$

if $|\mathbf{Q}| = s^N$. Altogether, we find, performing the integration with respect to r,

$$\left| \int_{\mathbf{Q}} u(\boldsymbol{x})\, d\boldsymbol{x} - |\mathbf{Q}| u(\mathbf{0}) \right| \le \frac{1}{1 - N/p} \|\nabla u\|_{L^p(\mathbf{Q};\mathbb{R}^N)} s^{1+N(p-1)/p}.$$

Dividing through $|\mathbf{Q}| = s^N$, we finally get

$$\left| \frac{1}{|\mathbf{Q}|} \int_{\mathbf{Q}} u(\boldsymbol{x})\, d\boldsymbol{x} - u(\mathbf{0}) \right| \le \frac{s^{1-N/p}}{1 - N/p} \|\nabla u\|_{L^p(\mathbf{Q};\mathbb{R}^N)}.$$

Step 3. Because in the previous step the cube $\mathbf{Q}$ was simply assumed to contain the origen $\mathbf{0}$, and the two sides on the last inequality are translation invariant, we can conclude that

$$\left| \frac{1}{|\mathbf{Q}|} \int_{\mathbf{Q}} u(\boldsymbol{x})\, d\boldsymbol{x} - u(\mathbf{y}) \right| \le \frac{s^{1-N/p}}{1 - N/p} \|\nabla u\|_{L^p(\mathbf{Q};\mathbb{R}^N)}$$

is also correct for any such cube $\mathbf{Q}$, and any $\mathbf{y} \in \mathbf{Q}$. In particular, if $\mathbf{y}$ and $\mathbf{z}$ are two given, arbitrary points, and $\mathbf{Q}$ is a cube containing them with side, say, $s = 2|\mathbf{y} - \mathbf{z}|$, then

$$\begin{aligned} |u(\mathbf{y}) - u(\mathbf{z})| \le & \left| \frac{1}{|\mathbf{Q}|} \int_{\mathbf{Q}} u(\boldsymbol{x})\, d\boldsymbol{x} - u(\mathbf{z}) \right| + \left| \frac{1}{|\mathbf{Q}|} \int_{\mathbf{Q}} u(\boldsymbol{x})\, d\boldsymbol{x} - u(\mathbf{y}) \right| \\ \le & C|\mathbf{y} - \mathbf{z}|^{1-N/p} \|\nabla u\|_{L^p(\mathbb{R}^N;\mathbb{R}^N)}, \end{aligned}$$

for a constant C depending on N and p but independent of u.

Theorem 9.5 *If $\Omega \subset \mathbb{R}^N$ is a domain, and $p > N$, every function $u \in W_0^{1,p}(\Omega)$*[1] *is such that*

$$|u(\mathbf{y}) - u(\mathbf{z})| \le C|\mathbf{y} - \mathbf{z}|^{1-N/p} \|\nabla u\|_{L^p(\Omega;\mathbb{R}^N)}$$

for every $\mathbf{y}, \mathbf{z} \in \Omega$, and a constant C independent of u. In particular, bounded sets of $W_0^{1,p}(\Omega)$ are equicontinuous, and

$$W_0^{1,p}(\Omega) \subset L^\infty(\Omega).$$

[1] This is understood in the sense that u can be redefined in a negligible set in such a way that the modified function complies with this statement.

9.3.2 The General Case

Note that for the particular case $\Omega = \mathbb{R}^N$, the two spaces $W^{1,p}(\mathbb{R}^N)$ and $W_0^{1,p}(\mathbb{R}^N)$ are identical, and so all of our results are valid for this special domain (without boundary). Recall Remark 7.3. In this section, we would like to explore the extension of the same inequalities we have examined in the previous section to functions u in $W^{1,p}(\Omega)$ without demanding to have a vanishing trace around $\partial\Omega$. This job typically restricts more the domains Ω where that goal can be achieved.

Since the treatment of the general situation is rather technical, though important, and relies on the extension property indicated in the Introduction of this chapter, we will not cover such. Instead, and as an example of how the inequality in Theorem 9.2 demands the explicit presence of the L^1-norm of the function u, we will redo the above calculations for a cube $\Omega = \mathbf{Q}$ in $\mathbb{R}^N$ with edges of size $|\mathbf{Q}|^{1/N}$.

We retake identity (9.6) to write

$$|u(\boldsymbol{x}', x)| \leq |u(\boldsymbol{x}', y)| + \int_y^x \left| \frac{\partial u}{\partial x_N}(\boldsymbol{x}', z) \right| dz. \tag{9.10}$$

The numbers x and y belong to $J_{\boldsymbol{x}'}$ where, according to (7.11),

$$\Omega = \pi_N \Omega \times J_{N,\boldsymbol{x}'} \mathbf{e}_N,$$

and the same can be done for every coordinate direction $\mathbf{e}_i$, $i = 1, 2, \dots N$.

If we integrate this inequality in the variable y in the two subsets

$$I_{\boldsymbol{x}'} = J_{\boldsymbol{x}'} \cap (\inf J_{\boldsymbol{x}'}, x), \qquad D_{\boldsymbol{x}'} = J_{\boldsymbol{x}'} \cap (x, \sup J_{\boldsymbol{x}'}),$$

we arrive at

$$|I_{\boldsymbol{x}'}|\, |u(\boldsymbol{x}', x)| \leq \int_{I_{\boldsymbol{x}'}} |u(\boldsymbol{x}', y)|\, dy + \int_{I_{\boldsymbol{x}'}} \int_y^x \left| \frac{\partial u}{\partial x_N}(\boldsymbol{x}', z) \right| dz\, dy,$$
$$|D_{\boldsymbol{x}'}|\, |u(\boldsymbol{x}', x)| \leq \int_{D_{\boldsymbol{x}'}} |u(\boldsymbol{x}', y)|\, dy + \int_{D_{\boldsymbol{x}'}} \int_x^y \left| \frac{\partial u}{\partial x_N}(\boldsymbol{x}', z) \right| dz\, dy.$$

If we use Fubini's theorem in the two double integrals (assuming the partial derivative extended by zero when necessary), we find that

$$|I_{\boldsymbol{x}'}|\, |u(\boldsymbol{x}', x)| \leq \int_{I_{\boldsymbol{x}'}} |u(\boldsymbol{x}', y)|\, dy + \int_{I_{\boldsymbol{x}'}} (z - \inf J_{\boldsymbol{x}'}) \left| \frac{\partial u}{\partial x_N}(\boldsymbol{x}', z) \right| dz,$$
$$|D_{\boldsymbol{x}'}|\, |u(\boldsymbol{x}', x)| \leq \int_{D_{\boldsymbol{x}'}} |u(\boldsymbol{x}', y)|\, dy + \int_{D_{\boldsymbol{x}'}} (\sup J_{\boldsymbol{x}'} - z) \left| \frac{\partial u}{\partial x_N}(\boldsymbol{x}', z) \right| dz,$$

or even

$$|I_{\boldsymbol{x}'}|\,|u(\boldsymbol{x}',x)| \le \int_{I_{\boldsymbol{x}'}} |u(\boldsymbol{x}',y)|\,dy + \int_{J_{\boldsymbol{x}'}} (z-\inf J_{\boldsymbol{x}'}) \left|\frac{\partial u}{\partial x_N}(\boldsymbol{x}',z)\right|\,dz,$$

$$|D_{\boldsymbol{x}'}|\,|u(\boldsymbol{x}',x)| \le \int_{D_{\boldsymbol{x}'}} |u(\boldsymbol{x}',y)|\,dy + \int_{J_{\boldsymbol{x}'}} (\sup J_{\boldsymbol{x}'}-z) \left|\frac{\partial u}{\partial x_N}(\boldsymbol{x}',z)\right|\,dz.$$

If we add up these two inequalities, we have

$$d(J_{\boldsymbol{x}'})|u(\boldsymbol{x}',x)| \le \int_{J_{\boldsymbol{x}'}} |u(\boldsymbol{x}',y)|\,dy + d(J_{\boldsymbol{x}'}) \int_{J_{\boldsymbol{x}'}} \left|\frac{\partial u}{\partial x_N}(\boldsymbol{x}',z)\right|\,dz,$$

where

$$d(J_{\boldsymbol{x}'}) = \sup J_{\boldsymbol{x}'} - \inf J_{\boldsymbol{x}'}$$

is the diameter of $J_{\boldsymbol{x}'}$. Dividing through by this positive number $d(J_{\boldsymbol{x}'})$, we are led to the inequality

$$|u(\boldsymbol{x}',x)| \le \frac{1}{d(J_{\boldsymbol{x}'})} \int_{J_{\boldsymbol{x}'}} |u(\boldsymbol{x}',z)|\,dz + \int_{J_{\boldsymbol{x}'}} \left|\frac{\partial u}{\partial x_N}(\boldsymbol{x}',z)\right|\,dz.$$

This is a possible replacement of inequality (9.7) to the general situation of a function u in $W^{1,1}(\Omega)$. Suppose, in particular, that $\Omega = \mathbf{Q}$ is a certain fixed cube. Then

$$|u(\boldsymbol{x}',x)| \le |\mathbf{Q}|^{-1/N} \int_{J_{\boldsymbol{x}'}} |u(\boldsymbol{x}',z)|\,dz + \int_{J_{\boldsymbol{x}'}} \left|\frac{\partial u}{\partial x_N}(\boldsymbol{x}',z)\right|\,dz.$$

If we now define

$$u_i(\boldsymbol{x}'_i) = |\mathbf{Q}|^{-1/N} \int_{J_{\boldsymbol{x}'_i}} |u(\boldsymbol{x}'_i,z)|\,dz + \int_{J_{\boldsymbol{x}'_i}} \left|\frac{\partial u}{\partial x_i}(\boldsymbol{x}'_i,z)\right|\,dz$$

for $i = 1, 2, \dots, N$, we have exactly the same inequality (9.8). Since

$$\|u_i\|_{L^1(\pi_i\mathbf{Q})} \le |\mathbf{Q}|^{-1/N}\|u\|_{L^1(\mathbf{Q})} + \|\nabla u\|_{L^1(\mathbf{Q};\mathbb{R}^N)},$$

the same use of Lemma 9.1 in this situation yields

$$\|u\|_{L^{N/(N-1)}(\mathbf{Q})} \le |\mathbf{Q}|^{-1/N}\|u\|_{L^1(\mathbf{Q})} + \|\nabla u\|_{L^1(\mathbf{Q};\mathbb{R}^N)}.$$

It does not seem possible to extend this inequality for more general domains than just cubes.

The following is a full summary of the fundamental results of this section.

Theorem 9.6 *Let $\Omega \subset \mathbb{R}^N$ be a C^1-domain (its boundary $\partial\Omega$ is a C^1-manifold of dimension $N-1$). Let $p \in [1, +\infty]$.*

1. *Subcritical case $1 \le p < N$:*

$$W^{1,p}(\Omega) \subset L^{p^*}(\Omega), \quad \frac{1}{p^*} = \frac{1}{p} - \frac{1}{N}.$$

2. *Critical case $p = N$:*

$$W^{1,p}(\Omega) \subset L^q(\Omega), \quad q \in [p, +\infty).$$

3. *Supercritical case $p > N$:*

$$W^{1,p}(\Omega) \subset L^\infty(\Omega).$$

Moreover,

1. *if Ω is bounded, and $p < N$, then*

$$W^{1,p}(\Omega) \subset L^q(\Omega), \quad q \in [1, p^*];$$

2. *all the above injections are continuous;*
3. *in the supercritical case $p > N$, we have the inequality*

$$|u(\boldsymbol{x}) - u(\boldsymbol{y})| \le C\|u\|_{W^{1,p}(\Omega)}|\boldsymbol{x} - \boldsymbol{y}|^{1-N/p} \tag{9.11}$$

valid for every $u \in W^{1,p}(\Omega)$ and every couple $\boldsymbol{x}, \boldsymbol{y} \in \Omega$, with a constant C depending only on Ω, p and dimension N; in particular, u admits a continuous representative in Ω.

It is also important to highlight how these results, together with the ideas in the proof of Proposition 7.10 also yield the following fine point.

Theorem 9.7 *For a bounded and C^1 domain $\Omega \subset \mathbb{R}^N$, the following are compact injections*

$$W^{1,p}(\Omega) \subset L^q(\Omega), \quad q \in [1, p^*],$$

$$W^{1,p}(\Omega) \subset L^q(\Omega), \quad q \in [1, +\infty),$$

$$W^{1,p}(\Omega) \subset C(\overline{\Omega}),$$

in the subcritical, critical and supercritical cases, respectively.

Proof The supercritical case is a direct consequence of (9.11) and the classical Arzelá-Ascoli theorem based on equicontinuity. Both the subcritical and critical cases, are a consequence of the same proof of Proposition 7.10. Note that this explains why the limit exponent p^* can be included in the subcritical case $p < N$, while it is not so for the critical case $p = N$. □

9.3.3 Higher-Order Sobolev Spaces

One can definitely apply Sobolev inequalities to first partial derivatives of functions in $W^{2,p}(\Omega)$, thus leading to better properties of functions in this space. We simply state here one main general result which does not deserve any further comment as it can be deduced inductively on the order of derivation.

Theorem 9.8 *Suppose $\Omega \subset \mathbb{R}^N$ is a C^1-, bounded domain. Let $u \in W^{k,p}(\Omega)$.*

1. *If $k < N/p$ and $1/q = 1/p - k/N$, then $u \in L^q(\Omega)$, and*

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{W^{k,p}(\Omega)}$$

   for a positive constant C independent of u, and depending on k, p, N and Ω.
2. *If $k > N/p$, put*

$$\alpha = \begin{cases} [N/p] + 1 - N/p, & N/p \text{ is not an integer}, \\ \text{any positive value } < 1, & N/p \text{ is an integer}. \end{cases}$$

   Then[2] *$u \in C^{k-[N/p]-1,\alpha}(\Omega)$, and*

$$|u^\gamma(\boldsymbol{x}) - u^\gamma(\boldsymbol{y})| \leq C\|u\|_{W^{k,p}(\Omega)}|\boldsymbol{x} - \boldsymbol{y}|^\alpha$$

   where u^γ is any derivative of u of order $k - [N/p] - 1$, $\boldsymbol{x}$, $\boldsymbol{y}$ are two arbitrary points in Ω, and the constant C only depends on k, p, N, α, and Ω.

9.4 Regularity of Domains, Extension, and Density

We have examined above two particular situations where functions in a Sobolev space Ω can be extended in an ordered way to functions in all of space $\mathbb{R}^N$. We would like to briefly describe, without entering into details, a more general method where this extension can be carried through. Basically, extensions of functions

[2] The notation $[r]$ indicates integer part of the number r.

defined in Ω to a bigger set $\tilde{\Omega}$, requires some regular transformation $\Phi : \tilde{\Omega} \setminus \Omega \to \Omega$ to define the extension through

$$U(\boldsymbol{x}) = u(\Phi(\boldsymbol{x})) \text{ for } \boldsymbol{x} \in \tilde{\Omega} \setminus \Omega.$$

Suppose Ω is a C^1-domain according to Definition 7.3 with ϕ its defining C^1-function, and $\epsilon > 0$ the strip parameter around $\partial\Omega$.

We learn from ODEs courses, that the flow

$$\Phi(t, \boldsymbol{x}) : [0, T) \times \mathbb{R}^N \to \mathbb{R}^N, \quad \Phi'(t, \boldsymbol{x}) = \nabla\phi(\Phi(t, \boldsymbol{x})), \Phi(0, \boldsymbol{x}) = \boldsymbol{x},$$

corresponding to the dynamical system

$$\boldsymbol{x}'(t) = \nabla\phi(\boldsymbol{x}(t)) \tag{9.12}$$

is a continuous mapping in the variable $\boldsymbol{x}$. Moreover, we also define the function $t(\boldsymbol{x})$ through the following condition

$$t(\boldsymbol{x}) : \{-\epsilon < \phi \le 0\} \to \{0 \le \phi < \epsilon\}, \quad \phi(\Phi(t(\boldsymbol{x}), \boldsymbol{x})) = -\phi(\boldsymbol{x}).$$

This function is as regular as the flow Φ, and, moreover

$$t(\boldsymbol{x}) = 0 \text{ for every } \boldsymbol{x} \in \partial\Omega.$$

Note that integral curves of (9.12) travel perpendicularly into Ω.

Though the following result can be treated in the generality of Definition 7.3, for the sake of simplicity and to avoid some additional technicalities, we will assume that the function ϕ is C^2 to rely on the C^1-feature of the flow mapping Φ.

Proposition 9.2 *Let Ω, ϕ, and ϵ be as in Definition 7.3 with ϕ, a C^2-function, and $\partial\Omega$, bounded. Let $u \in W^{1,p}(\Omega)$, and define*

$$U(\boldsymbol{x}) = \begin{cases} u(\boldsymbol{x}), & \boldsymbol{x} \in \Omega, \\ \left(1 - \frac{1}{\epsilon}\phi(\Phi(t(\boldsymbol{x}), \boldsymbol{x}))\right) u(\Phi(t(\boldsymbol{x}), \boldsymbol{x})), & \boldsymbol{x} \in \{-\epsilon < \phi < 0\}, \\ 0, & \boldsymbol{x} \in \mathbb{R}^N \setminus \Omega_\epsilon. \end{cases}$$

Then $U \in W^{1,p}(\mathbb{R}^N)$ is an extension of u, and

$$\|U\|_{W^{1,p}(\mathbb{R}^N)} \le C\|u\|_{W^{1,p}(\Omega)},$$

for some constant C depending on Ω through ϕ and ϵ.

With the help of the extension operation, it is easy to show the density of smooth functions in Sobolev spaces. The proof utilizes the ideas of Sect. 2.29 in a high-dimensional framework in the spirit of Corollary 7.1.

Corollary 9.1 *The restrictions of functions in $C^\infty(\mathbb{R}^N)$ with compact support make up a dense subspace of $W^{1,p}(\Omega)$ if Ω is C^1.*

9.5 An Existence Theorem Under More General Coercivity Conditions

Though more general examples could be dealt with, we will restrict attention to functionals of type (9.4)

$$I(u) = \int_\Omega \left(\frac{1}{2}|\nabla u(\boldsymbol{x})|^2 + f(u(\boldsymbol{x}))\right) d\boldsymbol{x} \tag{9.13}$$

to understand how, taking advantage of the better integrability properties that we have learnt in previous sections, one can relax the coercivity condition in our Theorem 8.2, and yet retain the existence of minimizers for the functional under typical Dirichlet boundary conditions. We therefore stick in this section to the model problem

$$\text{Minimize in } u(\boldsymbol{x}) \in H^1(\Omega): \quad \int_\Omega \left(\frac{1}{2}|\nabla u(\boldsymbol{x})|^2 + f(u(\boldsymbol{x}))\right) d\boldsymbol{x} \tag{9.14}$$

under

$$u - u_0 \in H_0^1(\Omega), \quad u_0 \in H^1(\Omega), \text{ given.}$$

Our goal is then to find more general conditions on function $f(u)$ that still allow to retain existence of minimizers. In particular, some of these conditions permit non-linearities $f(u)$ not bounded from below. In all cases, the main concern is to recover the necessary coercivity. The basic idea is to arrange things so that the contribution coming from the function f can be absorbed by the term involving the square of the gradient. Some of the techniques utilized are typical in this kind of calculations.

Theorem 9.9

1. *Suppose $f(u): \mathbb{R} \to \mathbb{R}$ is a continuous, bounded-from-below function. Then there are minimizers for (9.14).*
2. *Let $f(u)$ be such that*

$$|f(u)| \le \lambda u^2 + C, \quad 0 \le \lambda < \frac{1}{8C_P^2},$$

where $C_P > 0$ is the best constant for Poincaré's inequality in $H_0^1(\Omega)$. Then there are minimizers for (9.14).

3. *If the function $f(u)$ is such that*

$$|f(u)| \le c|u|^r + C, \quad r < 2, c > 0,$$

then there are minimizers for our problem.

Proof The first situation is immediate because the coercivity condition (8.6) for $p = 2$ is valid as soon as $f(u)$ is bounded from below by some constant $C \in \mathbb{R}$, i.e.

$$\frac{1}{2}|\mathbf{u}|^2 + C \le \frac{1}{2}|\mathbf{u}|^2 + f(u).$$

The existence of minimizers is then a direct consequence of Theorem 8.2.

The second possibility is also easy. Note that, by Poincaré's inequality applied to the difference $u - u_0 \in H_0^1(\Omega)$,

$$\begin{aligned}
\int_\Omega |f(u(\boldsymbol{x}))|\, d\boldsymbol{x} \le & \lambda\|u\|^2_{L^2(\Omega)} + C|\Omega| \\
\le & 2\lambda\|u - u_0\|^2_{L^2(\Omega)} + 2\lambda\|u_0\|^2_{L^2(\Omega)} + C|\Omega| \\
\le & 2\lambda C_P^2\|\nabla(u - u_0)\|^2_{L^2(\Omega;\mathbb{R}^N)} + 2\lambda\|u_0\|^2_{L^2(\Omega)} + C|\Omega| \\
\le & 4\lambda C_P^2\|\nabla u\|^2_{L^2(\Omega;\mathbb{R}^N)} + 4\lambda C_P^2\|\nabla u_0\|^2_{L^2(\Omega;\mathbb{R}^N)} \\
& + 2\lambda\|u_0\|^2_{L^2(\Omega)} + C|\Omega| \\
= & 4\lambda C_P^2\|\nabla u\|^2_{L^2(\Omega;\mathbb{R}^N)} + \tilde{C},
\end{aligned}$$

where $\tilde{C}$ is a constant independent of u. Hence,

$$\left(\frac{1}{2} - 4\lambda C_P^2\right)\|\nabla u\|^2_{L^2(\Omega;\mathbb{R}^N)} - \tilde{C} \le \int_\Omega \left(\frac{1}{2}|\nabla u(\boldsymbol{x})|^2 + f(u(\boldsymbol{x}))\right) d\boldsymbol{x},$$

with

$$\frac{1}{2} - 4\lambda C_P^2 > 0.$$

This is the coercivity required to recover existence of minimizers through the direct method.

For the third case, by Young's inequality Lemma 2.1, applied to the factors, for an arbitrary $\delta > 0$ and $\boldsymbol{x} \in \Omega$,

$$a = \delta|u(\boldsymbol{x})|^r, \quad b = \frac{1}{\delta},$$

with the exponents

$$p = \frac{2}{r} > 1, \quad q = \frac{r}{2-r},$$

we see that

$$|u(\boldsymbol{x})|^r = |ab| \leq \frac{1}{p}\delta^p |u(\boldsymbol{x})|^2 + \frac{1}{q}\frac{1}{\delta^q}.$$

An integration in Ω leads to

$$\int_\Omega |u(\boldsymbol{x})|^r \, d\boldsymbol{x} \leq \frac{\delta^p}{p}\|u\|^2_{L^2(\Omega)} + \frac{|\Omega|}{q\delta^q}.$$

From this inequality, and following along the calculations of the previous case, we find that

$$\int_\Omega \left(\frac{1}{2}|\nabla u(\boldsymbol{x})|^2 + f(u(\boldsymbol{x}))\right) d\boldsymbol{x} \geq \left(\frac{1}{2} - \frac{4c\delta^p}{p}C_P^2\right)\|\nabla u\|^2_{L^2(\Omega;\mathbb{R}^N)} + \tilde{C}(\delta),$$

where the constant $\tilde{C}(\delta)$ turns out to be

$$\tilde{C}(\delta) = -\frac{4c\delta^p}{p}C_P^2\|\nabla u_0\|^2_{L^2(\Omega;\mathbb{R}^N)} - \frac{2c\delta^p}{p}C_P^2\|u_0\|^2_{L^2(\Omega;\mathbb{R}^N)} - \frac{c|\Omega|}{q\delta^q} - C.$$

Since we still have the parameter δ to our disposal, it suffices to select it in such a way that

$$0 \leq \delta < \left(\frac{p}{8cC_P^2}\right)^{1/p},$$

to ensure the necessary coercivity, and to conclude the proof. □

9.6 Critical Point Theory

In this section, we become interested in solutions of the critical function equation (the Euler-Lagrange equation), not coming from minimization, for a functional of the same kind (9.13); that is to say, critical functions that are not minimizers. The emphasis is placed on the fundamental compactness property of Palais-Smale that permits to translate some of the intuitive ideas for functions of several (finite) variables to the infinite-dimensional setting. Unfortunately, checking details in a complete way requires quite a good deal of calculations with constants, exponents,

norms in different spaces, various easy estimates and inequalities with numbers and functions, etc. In particular, we will have an opportunity to see the relevance of Poincaré's and Sobolev inequalities. The basic relevant concepts are exactly as in the finite-dimensional case.

Definition 9.1 Let $E : \mathbb{H} \to \mathbb{R}$ be a C^1-functional defined in a Hilbert space $\mathbb{H}$.

1. A vector $\mathbf{u} \in \mathbb{H}$ is critical for E if $E'(\mathbf{u}) = \mathbf{0}$.
2. A number c is a critical value for E if there is $\mathbf{u} \in \mathbb{H}$ such that

$$E(\mathbf{u}) = c, \quad E'(\mathbf{u}) = \mathbf{0}.$$

The basic heuristic principle to detect critical points which are not minimizers is the following. When one is seeking a minimizer, one is led to minimize "in every possible direction or way". If, on the other hand, we look for critical points (where the derivative or gradient must vanish) of a function or functional that may not be minimizers, the first attempt would be to minimize "in all but one direction or dimension". This is the simplest version of a primitive min-max principle: minimize the maximum across a family of objects of a finite dimension. It can be made specific in the following intuitive way.

Let $E : \mathbb{H} \to \mathbb{R}$ be a C^1-functional over a Hilbert space $\mathbb{H}$. Take two vectors $\mathbf{u}_0$, $\mathbf{u}_1$ in $\mathbb{H}$, and consider the class Γ of smooth curves, regarded as one-dimensional objects,

$$\Gamma = \{\gamma(t) : [0, 1] \to \mathbb{H} : \gamma(0) = \mathbf{u}_0, \gamma(1) = \mathbf{u}_1\}.$$

For each such curve $\gamma \in \Gamma$, we would like to detect the maximum of the composition with E

$$\max_{t\in[0,1]} E(\gamma(t)),$$

which is obviously attained at some t_M that will, most likely, depend on γ. Then, we take the minimum (infimum) of all these maximum values across the set of curves in Γ

$$c = \inf_{\gamma\in\Gamma} \max_{t\in[0,1]} E(\gamma(t)),$$

and trust that this value ought to be critical. There are three main concerns:

1. c is a finite value (it suffices that E is bounded from below, though this is not always the case);
2. c is indeed a true critical value so that there some $\mathbf{u} \in \mathbb{H}$ with $E'(\mathbf{u}) = \mathbf{0}$ and $E(\mathbf{u}) = c$;
3. c is not a critical value for a minimizer.

The subtle issue is to ensure that c is a critical value. This would be almost immediate in a finite-dimensional setting. But it requires some caution in an infinite dimensional situation that, as in the case of weak convergence, is related to lack of compactness. The nature of the functional E needs to be strengthened, but we will see exactly in which step and what that reinforcement should be.

Imagine that a certain $c \in \mathbb{R}$ is not a critical value for the functional E. This means that the derivative cannot vanish in the region

$$\{c - \epsilon \leq E(\mathbf{u}) \leq c + \epsilon\} \equiv \{\mathbf{u} \in \mathbb{H} : E(\mathbf{u}) \in [c - \epsilon, c + \epsilon]\},$$

at least for some small positive ϵ. If we are working in finite dimension, this implies that the derivative E' is uniformly away from zero in such a region, i.e. there is some $\delta > 0$ (depending on ϵ), such that

$$\{c - \epsilon \leq E(\mathbf{u}) \leq c + \epsilon\} \subset \{\|E'\| \geq \delta\}.$$

However, in an infinite-dimensional setting, there might be a sequence $\{\mathbf{u}_j\}$ such that

$$E(\mathbf{u}_j) \in [c - \epsilon, c + \epsilon], \quad E'(\mathbf{u}_j) \to \mathbf{0}.$$

Such sequence could even be uniformly bounded in $\mathbb{H}$, but only converge in a weak sense to some $\mathbf{u}$. The point is that because the convergence is only weak, we loose all information concerning $E(\mathbf{u})$ and $E'(\mathbf{u})$. There is no way to ensure whether $\mathbf{u}$ is a critical point for E. This is precisely what needs to be adjusted for functionals defined in infinite-dimensional spaces.

Definition 9.2 A C^1-functional $E : \mathbb{H} \to \mathbb{R}$ defined on a Hilbert space $\mathbb{H}$ is said to enjoy the Palais-Smale property if for every sequence $\{\mathbf{u}_j\} \subset \mathbb{H}$ such that $\{E(\mathbf{u}_j)\}$ is a bounded sequence of numbers and $E'(\mathbf{u}_j) \to \mathbf{0}$ (strongly) in $\mathbb{H}$, we can always find some subsequence converging in $\mathbb{H}$. Sequences $\{\mathbf{u}_j\}$ with

$$|E(\mathbf{u}_j)| \leq M, \quad E'(\mathbf{u}_j) \to \mathbf{0},$$

for some fixed M, are called Palais-Smale sequences (for E).

Before proceeding, since this concept may sound a bit artificial to readers, it is interesting to isolate the main family of functionals that comply with the Palais-Smale condition. Recall that $\mathbf{1} : \mathbb{H} \to \mathbb{H}$ is the identity operator.

Proposition 9.3 *Suppose that Palais-Smale sequences are uniformly bounded (for instance if E is coercive), for a C^1-functional $E : \mathbb{H} \to \mathbb{R}$, and*

$$E' = \mathbf{1} + \mathbf{K}, \quad \mathbf{K} : \mathbb{H} \to \mathbb{H}, \ \textit{a compact operator}.$$

Then E enjoys the Palais-Smale property.

Proof If $\{\mathbf{u}_j\}$ is a Palais-Smale sequence for E, it is bounded, and, for some subsequence, not relabeled, we would have the weak convergence $\mathbf{u}_j \rightharpoonup \mathbf{u}$. The compactness of $\mathbf{K}$ would lead to a further subsequence, if necessary, with $\mathbf{Ku}_j \to \tilde{\mathbf{u}}$. Thus

$$\mathbf{u}_j = E'(\mathbf{u}_j) - \mathbf{Ku}_j \to -\tilde{\mathbf{u}}$$

strongly. □

We now focus on the most studied situation in which we take $\mathbb{H} = H_0^1(\Omega)$ for a bounded domain $\Omega \subset \mathbb{R}^N$, and $E : \mathbb{H} \to \mathbb{R}$ defined by

$$E(u) = \int_\Omega \left(\frac{1}{2}|\nabla u(\boldsymbol{x})|^2 + f(u(\boldsymbol{x})) \right) d\boldsymbol{x} \tag{9.15}$$

with f a C^1-function. Most of the statements of the technical assumptions for the non-linearity $f(u)$ would require the distinction between dimension $N = 1, 2$ and $N > 2$. As a rule, we will assume in what follows that $N > 2$, and will leave the cases $N = 1, 2$ for the exercises.

Lemma 9.3

1. *If*

$$|f'(u)| \le c|u|^r + C, \quad c > 0, r < \frac{N+2}{N-2}, \tag{9.16}$$

then the functional E in (9.15) *is C^1, and $E' = \mathbf{1} + \mathbf{K}$ with $\mathbf{K}$, compact.*

2. *If, in addition to* (9.16),

$$|f(u)| \le \lambda |u|^2 + C,$$

with

$$0 \le \lambda < \frac{1}{8C_P^2},$$

where C_P is the best constant in Poincaré's inequality in $H_0^1(\Omega)$, then Palais-Smale sequences are uniformly bounded in $H_0^1(\Omega)$, and by Proposition 9.3, functional E in (9.15) *complies with the Palais-Smale condition.*

3. *If, in addition to* (9.16),

$$f(u) - \lambda f'(u)u \ge C, \quad \lambda < \frac{1}{2}, \tag{9.17}$$

then Palais-Smale sequences are uniformly bounded in $H_0^1(\Omega)$ *(though this time the funcional might not be coercive), and again by Proposition 9.3, functional* E *in* (9.15) *complies with the Palais-Smale condition.*

Proof In order to apply Proposition 9.3 to functional (9.15), we need to check three things:

1. E is C^1;
2. $E' = \mathbf{1} + \mathbf{K}$ with $\mathbf{K}$, compact;
3. Palais-Smale sequences are uniformly bounded in $\mathbb{H}$.

For the first point, and according to Definitions 2.13 and 2.14, we need to calculate the directional derivative

$$\left.\frac{d}{d\epsilon} I(u + \epsilon U)\right|_{\epsilon=0}$$

and make sure that it is a continuous operation in u and linear in U. Based on the differentiability of f, it is easy to conclude that

$$\left.\frac{d}{d\epsilon}\int_\Omega \left(\frac{1}{2}|\nabla u(\boldsymbol{x}) + \epsilon\nabla U(\boldsymbol{x})|^2 + f(u(\boldsymbol{x}) + \epsilon U(\boldsymbol{x}))\right) d\boldsymbol{x}\right|_{\epsilon=0}$$

is given by the expression

$$\int_\Omega \left(\nabla u(\boldsymbol{x}) \cdot \nabla U(\boldsymbol{x}) + f'(u(\boldsymbol{x}))U(\boldsymbol{x})\right) d\boldsymbol{x},$$

that indeed it is continuous in u, since f' is, and linear in U. Moreover, by Lemma 2.7, the derivative can be identified as the unique minimizer of the quadratic variational problem that consists in minimizing in $H_0^1(\Omega)$ the functional

$$\int_\Omega \left[\frac{1}{2}|\nabla U(\boldsymbol{x})|^2 - \nabla u(\boldsymbol{x}) \cdot \nabla U(\boldsymbol{x}) - f'(u(\boldsymbol{x}))U(\boldsymbol{x})\right] d\boldsymbol{x}.$$

The unique solution U of this problem is furnished, for instance, by the versions of the Lax-Milgram lemma in Sect. 8.2, and can be identified with the unique solution $U(= I'(u))$ (for given u) of the linear PDE problem

$$-\operatorname{div}(\nabla U - \nabla u) - f'(u) = 0 \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega.$$

We clearly see that $U = I'(u) = u$ plus the operation $\mathbf{K}$ taking u into the solution $v \in H_0^1(\Omega)$ of the problem

$$-\Delta v = f'(u) \text{ in } \Omega, \quad v = 0 \text{ on } \partial\Omega.$$

By writing its weak formulation, using as a test function v itself, we see that

$$\int_\Omega |\nabla v(\boldsymbol{x})|^2 \, d\boldsymbol{x} = \int_\Omega f'(u(\boldsymbol{x}))v(\boldsymbol{x}) \, d\boldsymbol{x}.$$

Using Hölder's inequality on the integral on the right-hand side for exponents

$$p = \frac{2N}{N-2}, \quad q = \frac{2N}{N+2},$$

we arrive at

$$\|\nabla v\|^2_{L^2(\Omega)} \le \|v\|_{L^{2N/(N-2)}(\Omega)} \, \|f'(u)\|_{L^{2N/(N+2)}(\Omega)}.$$

The corresponding embedding inequality leaves us with

$$\|\nabla v\|_{L^2(\Omega)} \le \|f'(u)\|_{L^{2N/(N+2)}(\Omega)},$$

and the bound on the size of $f'(u)$ leads us to conclude that

$$\|\nabla v\|_{L^2(\Omega)} \le c\|u\|_{L^{(N+2)/(N-2)}(\Omega)} + C$$

for new constants $c > 0$ and C. Due to the compact embedding of $H^1_0(\Omega)$ into $L^2(\Omega)$, we see that indeed $\mathbf{K}$ is compact.

In case the upper-bounds for f' and f can be established in relationship to Poincaré's inequality, then the coercivity shown in the proof of the second part of Theorem 9.9 furnishes the final ingredient: Palais-Smale sequences are uniformly bounded.

For the last item, suppose $\{u_j\}$ is a Palais-Smale sequence

$$E(u_j) \le M, \quad E'(u_j) \to \mathbf{0},$$

that is

$$\int_\Omega \left(\frac{1}{2}|\nabla u_j(\boldsymbol{x})|^2 + f(u_j(\boldsymbol{x}))\right) d\boldsymbol{x} \le M,$$

$$\int_\Omega \left(\nabla u_j(\boldsymbol{x}) \cdot \nabla U(\boldsymbol{x}) + f'(u_j(\boldsymbol{x}))U(\boldsymbol{x})\right) d\boldsymbol{x} \to 0,$$

uniformly on bounded sets of test functions U's in $H^1_0(\Omega)$. In particular, for

$$U = \frac{1}{\|\nabla u_j\|_{L^2(\Omega)}} u_j,$$

we can deduce that

$$\int_\Omega \left(|\nabla u_j(\boldsymbol{x})|^2 + f'(u_j(\boldsymbol{x}))u_j(\boldsymbol{x}) \right) d\boldsymbol{x} \le \|\nabla u_j\|_{L^2(\Omega)}$$

for j sufficiently large. From all this information, we conclude that for the combination

$$Q \equiv \int_\Omega \left(\frac{1}{2}|\nabla u_j(\boldsymbol{x})|^2 + f(u_j(\boldsymbol{x})) \right) d\boldsymbol{x} \\ - \lambda \int_\Omega \left(|\nabla u_j(\boldsymbol{x})|^2 + f'(u_j(\boldsymbol{x}))u_j(\boldsymbol{x}) \right) d\boldsymbol{x}$$

we find, through (9.17),

$$(\frac{1}{2} - \lambda)\|\nabla u_j\|^2_{L^2(\Omega)} + C \le Q \le M + \lambda \|\nabla u_j\|_{L^2(\Omega)}.$$

Since the coefficient in front of the square is strictly positive, this inequality implies the uniform boundedness of the sequence of numbers $\{\|\nabla u_j\|\}$, as claimed. □

We finally need to face the min-max principle announced above. It is universally known as the mountain-pass lemma, as this term intuitively describes the situation. Recall the definition of the class of paths Γ when the two vectors $\mathbf{u}_0$ and $\mathbf{u}_1$ are given.

Theorem 9.10 *Let $\mathbb{H}$ be a Hilbert space and E, a C^1-functional defined on $\mathbb{H}$ that satisfies the Palais-Smale condition. If there are $\mathbf{u}_0$, $\mathbf{u}_1$ in $\mathbb{H}$, and*

$$0 < r < \|\mathbf{u}_0 - \mathbf{u}_1\|$$

such that

$$\max\{E(\mathbf{u}_0), E(\mathbf{u}_1)\} < m_r \equiv \inf_{\|\mathbf{u}-\mathbf{u}_0\|=r} E(\mathbf{u}),$$

then the value

$$c = \inf_{\gamma\in\Gamma} \max_{t\in[0,1]} E(\gamma(t)) \ge m_r$$

is a critical value for E.

Proof Since each element γ in Γ will intersect at some point the set where m_r is the infimum, because $r < \|\mathbf{u}_0 - \mathbf{u}_1\|$, we clearly see that $c \ge m_r$, and, hence, c cannot be a critical value corresponding to a (local) minimum.

We will proceed by contradiction assuming that c is not a critical value. If this is so, and E verifies the Palais-Smale condition, we claim that there are

$$\epsilon < \frac{1}{2}(m_r - \max\{E(\mathbf{u}_0), E(\mathbf{u}_1)\}), \quad \delta > 0,$$

with the property

$$E(\mathbf{u}) \in [c - \epsilon, c + \epsilon] \text{ implies } \|E'(\mathbf{u})\| \geq \delta. \tag{9.18}$$

Suppose this were not the case, so that we could find $\mathbf{u}_j$ with $E(\mathbf{u}_j) - c \to 0$, and yet $E'(\mathbf{u}_j) \to \mathbf{0}$. These conditions exactly mean that $\{\mathbf{u}_j\}$ would be a Palais-Smale sequence, and therefore, we would be able to find an accumulation vector $\mathbf{u}$ which, by continuity of E and E', would be a critical point at level c. If this situation is impossible, we could certainly find some such ϵ and δ for which (9.18) is correct.

By definition of c, there should be a path $\gamma \in \Gamma$ such that

$$E(\gamma(t)) \leq c + \epsilon, \quad t \in [0, 1].$$

The idea is to use the flow of E to produce a new feasible path $\tilde{\gamma}$ such that

$$E(\tilde{\gamma}(t)) \leq c - \epsilon.$$

This would be the contradiction, since again the definition of c makes this impossible. But calculations need to be performed quantitatively in a very precise way.

For each fixed $t \in [0, 1]$ for which

$$E(\gamma(t)) > c - \epsilon,$$

consider the infinite-dimensional gradient system

$$\sigma'(s) = -E'(\sigma(s)), \quad \sigma(0) = \gamma(t).$$

We will assume that E' is locally Lipschitz continuous to avoid some more technicalities. This requirement is not necessary, but if E' does not comply with this local lipschitzianity condition one needs to modify it appropriately. By Lemma 2.6, the previous gradient system is defined for all s positive, and its solution $\sigma(s; \gamma(t))$ depends continuously on the initial datum $\gamma(t)$. Furthermore, again by Lemma 2.6,

for $r > 0$,

$$E(\sigma(r)) - E(\gamma(t)) = \int_0^r \langle E'(\sigma(s)), \sigma'(s)\rangle \, ds$$
$$= - \int_0^r \|E'(\sigma(s))\|^2 \, ds$$
$$\leq - r\delta^2,$$

while $\sigma(s)$ stays in the region $\{E \in [c - \epsilon, c + \epsilon]\}$. Thus, while this is correct,

$$E(\sigma(r)) \leq E(\gamma(t)) - r\delta^2 \leq c - \epsilon, \tag{9.19}$$

as soon as

$$\frac{E(\gamma(t)) - \epsilon - c}{\delta^2} \leq r.$$

If we let $r(t)$ be the (continuous) minimum of the these values for which (9.19) holds, we find a new continuous path

$$\tilde{\gamma}(t) = \sigma(r(t); \gamma(t))$$

with

$$E(\tilde{\gamma}(t)) \leq c - \epsilon, \quad t \in [0, 1].$$

This contradicts the definition of c, and consequently, it must be a critical value. □

One of the principal applications of the results above which justifies the central role of the Palais-Smale condition for the critical point theory in infinite dimension refers again to the situation around (9.15) for which

$$E(u) : H_0^1(\Omega) \to \mathbb{R}, \quad E(u) = \int_\Omega \left(\frac{1}{2} |\nabla u(\boldsymbol{x})|^2 - f(u(\boldsymbol{x})) \right) d\boldsymbol{x}.$$

Note the change of sign in front of $f(u)$.

Theorem 9.11 *Consider this last problem with*

$$|f'(u)| \leq c|u|^r + C, \quad c > 0, 1 < r < \frac{N+2}{N-2}, \tag{9.20}$$

$$\lim_{u \to 0} \frac{f(u)}{u^2} = 0, \tag{9.21}$$

$$0 < f(u) \le \lambda f'(u)u, \quad 0 < \lambda < \frac{1}{2}, |u| \ge R, \tag{9.22}$$

for some positive R. Then there is some non-trivial critical function for the problem.

Proof Since condition (9.22) clearly implies (9.16), and, thus, by Proposition 9.3 and Lemma 9.3, the functional enjoys the Palais-Smale property, all that is left to do, to be able to apply Theorem 9.10, is to detect the two functions u_0, u_1 for which the situation in that theorem holds. We will show that $u_0 \equiv 0$ with $E(0) = 0$ is a strict local minimizer, and for every $v \in H_0^1(\Omega)$, there is always t with $E(tv) \le 0$. Hence if we take $u_1 = tv$, the condition of the trivial function being a strict local minimum clearly indicates that we are in the situation of Theorem 9.10, and we can conclude our result by Theorem 9.10.

It is clear that $E(0) = 0$. We want to show that for some positive ρ, $E(u) > 0$, whenever $0 < \|u\| \le \rho$. Recall that

$$\|u\| = \|\nabla u\|_{L^2(\Omega)}.$$

Let $\epsilon > 0$ be arbitrary. From (9.21), there is some $\delta = \delta(\epsilon) > 0$ with the property

$$|f(u)| \le \frac{\epsilon}{2}u^2, \quad |u| \le \delta.$$

On the other hand, from (9.20) by integration, we find, for some constant $C = C(\epsilon)$,

$$|f(u)| \le C|u|^{r+1}, \quad |u| \ge \delta.$$

Altogether, we have

$$|f(u)| \le \frac{\epsilon}{2}u^2 + C|u|^{r+1}$$

for all u. We bear in mind this estimate, and go to estimating $E(u)$. Indeed,

$$E(u) \ge \frac{1}{2}\|\nabla u\|^2_{L^2(\Omega)} - \frac{\epsilon}{2}\|u\|^2_{L^2(\Omega)} - C\|u\|^{r+1}_{L^{r+1}(\Omega)}.$$

We now invoke two facts. First, Poincaré's inequality

$$\|u\|_{L^2(\Omega)} \le C_P\|\nabla u\|_{L^2(\Omega)};$$

secondly, the Sobolev inequality

$$\|u\|_{L^{r+1}(\Omega)} \le C_S\|\nabla u\|_{L^2(\Omega)}$$

for some constant. Note that $r + 1 < (2N)/(N - 2)$. We can hence estimate

$$E(u) \geq \left(\frac{1}{2} - \frac{\epsilon}{2}C_P^2\right) \|\nabla u\|^2_{L^2(\Omega)} - CC_S^{r+1} \|\nabla u\|^{r+1}_{L^2(\Omega)}.$$

If we take ϵ sufficiently small so that the coefficient in front on $\|u\|^2$ is positive, and realizing that $r + 1 > 2$, we can certainly conclude that $E(u) > 0$ for all non-trivial u in a certain ball around the trivial function. This is exactly what is meant by saying that $u \equiv 0$ is a strict local minimum.

Finally, it is elementary to check (Exercise 4) that (9.22) implies

$$f(u) \geq a|u|^p + b, \quad a > 0, b \in \mathbb{R}, p = \frac{1}{\lambda} > 2.$$

By using this estimate in our functional, we arrive at the upper bound

$$E(u) \leq \frac{1}{2}\|\nabla u\|^2_{L^2(\Omega)} - a\|u\|^p_{L^p(\Omega)} - b|\Omega|.$$

But replacing tu by u and letting t move, we find that the right-hand side converges to $-\infty$, because $p > 2$, as $t \to \infty$. Consequently, the same is correct for $E(tu)$. In particular, for each u given, there is some t sufficiently large so that $E(tu) < 0$. This was the other necessary ingredient. □

We will see more specific examples in the exercises.

9.7 Regularity. Strong Solutions for PDEs

We start with an extension of the identity (8.13).

Lemma 9.4 *Let $\Omega \subset \mathbb{R}^N$ be a regular, smooth domain with boundary $\partial\Omega$ which is a C^2-smooth, $(N - 1)$-dimensional manifold, possibly with several connected components, and unit, outer normal direction $\mathbf{n}(\boldsymbol{x})$.*

1. *If $u \in H^2(\Omega) \cap H^1_0(\Omega)$, then*

$$\int_\Omega |\nabla^2 u(\boldsymbol{x})|^2 \, d\boldsymbol{x} = \int_\Omega |\Delta u(\boldsymbol{x})|^2 \, d\boldsymbol{x} + \int_{\partial\Omega} H(\boldsymbol{x})|\nabla u(\boldsymbol{x}) \cdot \mathbf{n}(\boldsymbol{x})|^2 \, dS(\boldsymbol{x}),$$

where $H(\boldsymbol{x})$ is the curvature of $\partial\Omega$ at $\boldsymbol{x}$.

2. *If $u \in H^2(\Omega)$ with*

$$\nabla u(\boldsymbol{x}) \cdot \mathbf{n}(\boldsymbol{x}) = 0$$

then

$$\int_\Omega |\nabla^2 u(\boldsymbol{x})|^2 \, d\boldsymbol{x} = \int_\Omega |\Delta u(\boldsymbol{x})|^2 \, d\boldsymbol{x} + \int_{\partial\Omega} \nabla \left(\frac{1}{2} |\nabla u|^2 \right) \cdot \mathbf{n} \, dS(\boldsymbol{x}).$$

3. *In particular, if* $u \in H_0^2(\Omega)$, *then*

$$\int_\Omega |\nabla^2 u(\boldsymbol{x})|^2 \, d\boldsymbol{x} = \int_\Omega |\Delta u(\boldsymbol{x})|^2 \, d\boldsymbol{x}.$$

Proof The main computation involves a typical integration by parts. Proceeding by density, just as in Definition 7.6, suppose for the time being that u is C^∞- in Ω. Then

$$\begin{aligned}\int_\Omega |\nabla^2 u(\boldsymbol{x})|^2 \, d\boldsymbol{x} &= \int_\Omega \operatorname{tr} \left(\nabla^2 u(\boldsymbol{x}) \nabla^2 u(\boldsymbol{x}) \right) \, d\boldsymbol{x} \\ &= \sum_i \int_\Omega \nabla \left(\frac{\partial u}{\partial x_i} \right) \cdot \frac{\partial}{\partial x_i} (\nabla u) \, d\boldsymbol{x}.\end{aligned}$$

We can use the divergence theorem for smooth functions and domains twice in each of the terms of the last sum to write, keeping track of boundary contributions and putting $\mathbf{n} = (n_i)$,

$$\begin{aligned}\int_\Omega \nabla \left(\frac{\partial u}{\partial x_i} \right) \cdot \frac{\partial}{\partial x_i} (\nabla u) \, d\boldsymbol{x} &= - \int_\Omega \frac{\partial u}{\partial x_i} \frac{\partial}{\partial x_i} (\Delta u) \, d\boldsymbol{x} + \int_{\partial\Omega} \frac{\partial u}{\partial x_i} \frac{\partial}{\partial x_i} \nabla u \cdot \mathbf{n} \, dS(\boldsymbol{x}) \\ &= \int_\Omega \frac{\partial^2 u}{\partial x_i^2} \Delta u \, d\boldsymbol{x} - \int_{\partial\Omega} \frac{\partial u}{\partial x_i} n_i \Delta u \, dS(\boldsymbol{x}) \\ &\quad + \int_{\partial\Omega} \frac{\partial u}{\partial x_i} \frac{\partial}{\partial x_i} \nabla u \cdot \mathbf{n} \, dS(\boldsymbol{x}).\end{aligned}$$

The sum in the index i carries us, through the first identity above, to

$$\begin{aligned}\int_\Omega |\nabla^2 u(\boldsymbol{x})|^2 \, d\boldsymbol{x} &= \int_\Omega |\Delta u(\boldsymbol{x})|^2 \, d\boldsymbol{x} \\ &\quad + \int_{\partial\Omega} \left(\nabla u \, \nabla^2 u \, \mathbf{n} - \nabla u \cdot \mathbf{n} \, \Delta u \right) \, dS(\boldsymbol{x}).\end{aligned} \tag{9.23}$$

If $u = 0$ on $\partial\Omega$, then

$$|\nabla u| \mathbf{n} = \nabla u, \quad \nabla u \cdot \mathbf{n} = |\nabla u|,$$

on $\partial\Omega$ because this boundary becomes a part of the level set $\{u = 0\}$, and then

$$\int_\Omega |\nabla^2 u(\boldsymbol{x})|^2\,d\boldsymbol{x} = \int_\Omega |\Delta u(\boldsymbol{x})|^2\,d\boldsymbol{x} + \int_{\partial\Omega} |\nabla u| \left(\mathbf{n}\,\nabla^2 u\,\mathbf{n} - \Delta u\right)\,dS(\boldsymbol{x}).$$

But under our smoothness assumptions (Exercise 1 below)

$$\mathbf{n}\,\nabla^2 u\,\mathbf{n} - \Delta u = -\,\Delta|_{\partial\Omega}\,u + H\nabla u \cdot \mathbf{n} \tag{9.24}$$

at $\partial\Omega$ where

$$\Delta|_{\partial\Omega}$$

is the Laplace operator over $\partial\Omega$. If $u = 0$ on $\partial\Omega$, then

$$\Delta|_{\partial\Omega}\,u = 0,$$

and this yields our first formula.

For the second situation, simply note that by retaking formula (9.23), the second term in the surface integral vanishes, and we find

$$\int_\Omega |\nabla^2 u(\boldsymbol{x})|^2\,d\boldsymbol{x} = \int_\Omega |\Delta u(\boldsymbol{x})|^2\,d\boldsymbol{x} + \int_{\partial\Omega} \nabla\left(\frac{1}{2}|\nabla u|^2\right)\cdot \mathbf{n}\,dS(\boldsymbol{x}).$$

In both cases, a standard density argument yields the claimed formulas for functions in the respective Sobolev spaces. □

For $f(\boldsymbol{x}) \in L^2(\Omega)$, with Ω a C^2-domain as in the previous lemma, we would like to consider the second-order problem

$$\text{Minimize in } u(\boldsymbol{x}) \in H^2(\Omega)\cap H^1_0(\Omega): \quad \int_\Omega \left(|\nabla^2 u(\boldsymbol{x})|^2 + 2f(\boldsymbol{x})\Delta u(\boldsymbol{x})\right)\,d\boldsymbol{x}.$$

This is a standard second-order variational problem that does not require any special consideration.

Proposition 9.4 *The previous variational problem admits a unique minimizer $\tilde{u} \in H^2(\Omega)\cap H^1_0(\Omega)$.*

Proof The functional is well-defined; it is coercive in its feasible set, and strictly convex. Theorem 8.5 applies. Note that indeed it is a quadratic, strictly convex functional on its admissible set. □

The point is that, due to Lemma 9.4, our second-order, variational problem above is exactly the same, except for the constant term

$$\int_\Omega |f(\boldsymbol{x})|^2\,d\boldsymbol{x},$$

as

$$\text{Minimize in } u(\boldsymbol{x}) \in H^2(\Omega)\cap H_0^1(\Omega): \quad \int_\Omega |\Delta u(\boldsymbol{x}) + f(\boldsymbol{x})|^2\,d\boldsymbol{x} + \int_{\partial\Omega} H(\boldsymbol{x})|\nabla u(\boldsymbol{x})\cdot\mathbf{n}(\boldsymbol{x})|^2\,dS(\boldsymbol{x}).$$

This variational problem is a bit special in that the functional incorporates a term which is a surface integral, and we have not explored how to deal with such. Some examples have been considered in connection with Neumman boundary conditions. Yet, because of the equivalence of these two variational problems, as indicated, we do know that the minimizer $\tilde{u}$ in Proposition 9.4 must also be a minimizer for the same variational problem in the second form. In particular, the surface integral

$$\int_{\partial\Omega} H(\boldsymbol{x})|\nabla\tilde{u}(\boldsymbol{x})\cdot\mathbf{n}(\boldsymbol{x})|^2\,dS(\boldsymbol{x}) \tag{9.25}$$

is well-defined and is finite. If we now look at the restricted variational problem

$$\text{Minimize in } u \in H_0^2(\Omega): \quad \int_\Omega |\Delta\tilde{u}(\boldsymbol{x}) + \Delta u(\boldsymbol{x}) + f(\boldsymbol{x})|^2\,d\boldsymbol{x},$$

because $\tilde{u}+u$, for a feasible $u \in H_0^2(\Omega)$, does not change the boundary value of $\tilde{u}$ or of its normal derivative involved in the surface integral (1.2), we conclude that the trivial function $u \equiv 0$ is a minimizer of this last problem. As such, we can examine optimality conditions as derived in Sect. 8.8, to argue that

$$\int_\Omega (\Delta\tilde{u}(\boldsymbol{x}) + f(\boldsymbol{x}))\Delta v(\boldsymbol{x})\,d\boldsymbol{x} = 0 \tag{9.26}$$

for every feasible $v \in H_0^2(\Omega)$. Since we know that the classical problem

$$\Delta v = g \text{ in } \mathbf{B}, \qquad v = 0 \text{ on } \partial\mathbf{B}$$

for arbitrary smooth functions g and any ball $\mathbf{B}$ has a unique smooth solution v, we realize that (9.26) implies that indeed

$$\Delta\tilde{u}(\boldsymbol{x}) + f(\boldsymbol{x}) = 0 \tag{9.27}$$

for a.e. $\boldsymbol{x} \in \Omega$. This is essentially a consequence of Lemma 7.4.

Definition 9.3 We say that a function $u \in H^2(\Omega)$ is a strong solution of $\Delta u + f = 0$ for a function $f \in L^2(\Omega)$ if

$$\Delta u(\boldsymbol{x}) + f(\boldsymbol{x}) = 0$$

for a.e. $\boldsymbol{x} \in \Omega$.

Note how different the definition of weak and strong solutions for the same PDE are. This definition can, of course, be generalized to many more families of PDEs.

Our arguments above (9.27) show that $\tilde{u}$ is a strong solution of $\Delta u + f = 0$, and, hence, it is also the weak solution of the problem.

Theorem 9.12 *Let $u \in H_0^1(\Omega)$ be the unique minimizer of the problem*

$$\textit{Minimize in } v \in H_0^1(\Omega): \quad \int_\Omega \left[\frac{1}{2}|\nabla v(\boldsymbol{x})|^2 - f(\boldsymbol{x})v(\boldsymbol{x})\right] d\boldsymbol{x}$$

where $f \in L^2(\Omega)$, and Ω is a C^2-domain. Then $u \in H^2(\Omega)$ and $\Delta u + f = 0$ a.e. in Ω.

Proof Under the smoothness assumptions on Ω, the function $\tilde{u} \in H_0^1(\Omega)$ is a strong solution, i.e.

$$\Delta \tilde{u}(\boldsymbol{x}) + f(\boldsymbol{x}) = 0$$

for a.e. $\boldsymbol{x} \in \Omega$. Therefore for an arbitrary $v \in H_0^1(\Omega)$,

$$\int_\Omega (\Delta \tilde{u}(\boldsymbol{x}) + f(\boldsymbol{x}))v(\boldsymbol{x})\, d\boldsymbol{x} = 0.$$

But the integration-by-parts formula applied to the first term yields

$$\int_\Omega [\nabla \tilde{u}(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) - f(\boldsymbol{x})v(\boldsymbol{x})]\, d\boldsymbol{x} = 0.$$

The arbitrariness of $v \in H_0^1(\Omega)$ in this equality implies that $\tilde{u}$ must be the weak solutions of our problem, which belongs, then, to $H^2(\Omega)$. □

This result is the basis of the regularity theory for PDEs: under suitable smoothness assumptions on Ω and the right-hand side $f \in L^2(\Omega)$ of the PDE, the unique weak solution $u \in H_0^1(\Omega)$ turns out to belong to $H^2(\Omega)$. From here one can translate further regularity on f and on Ω into more regularity for the weak solution u. The smoothness of $\partial\Omega$ is unavoidable.

9.8 Eigenvalues and Eigenfunctions

Possibly, the most important case and application of Theorems 6.4 and 6.3 (see also Example 6.6) is that of the eigenvalues and eigenfunctions of the Laplace operator under vanishing Dirichlet boundary conditions. Consider the operator, for a given bounded domain $\Omega \subset \mathbb{R}^N$,

$$\mathbf{T} : L^2(\Omega) \mapsto L^2(\Omega), \quad U = \mathbf{T}u, -\Delta U = u \text{ in } \Omega, U \in H_0^1(\Omega).$$

In other words, for each $u \in L^2(\Omega)$, U is the unique minimizer in $H_0^1(\Omega)$ of the functional

$$\int_\Omega \left[\frac{1}{2}|\nabla U(\boldsymbol{x})|^2 - U(\boldsymbol{x})u(\boldsymbol{x})\right] d\boldsymbol{x}.$$

In particular, we have that

$$\int_\Omega [\nabla U(\boldsymbol{x}) \cdot \nabla V(\boldsymbol{x}) - u(\boldsymbol{x})V(\boldsymbol{x})] \, d\boldsymbol{x} = 0 \tag{9.28}$$

for every $V \in H_0^1(\Omega)$.

The operator $\mathbf{T}$ is compact and self-adjoint:

1. self-adjointness: for $u, v \in L^2(\Omega)$ we easily realize, if we set $U = \mathbf{T}u$, $V = \mathbf{T}v$, that

$$\begin{aligned}\langle u, V\rangle =& \int_\Omega u(\boldsymbol{x})V(\boldsymbol{x})\, d\boldsymbol{x}\\ =& \int_\Omega \nabla U(\boldsymbol{x}) \cdot \nabla V(\boldsymbol{x})\, d\boldsymbol{x}\\ =& \int_\Omega U(\boldsymbol{x})v(\boldsymbol{x})\, d\boldsymbol{x}\\ =&\langle U, v\rangle,\end{aligned}$$

   by (9.28);
2. compactness: this is a direct consequence of the fact that the operation

$$u \in L^2(\Omega) \mapsto U \in H_0^1(\Omega), \quad -\Delta U = u \text{ in } \Omega$$

   is continuous, because by (9.28) for $V = U$ itself,

$$\|U\|^2_{H^1(\Omega)} \le \|uU\|_{L^1(\Omega)} \le \|u\|_{L^2(\Omega)} \|U\|_{L^2(\Omega)},$$

and by Poincaré's inequality

$$\|U\|^2_{H^1(\Omega)} \le C\|u\|_{L^2(\Omega)} \|U\|_{H^1(\Omega)};$$

the compactness of the injection $H^1_0(\Omega)$ into $L^2(\Omega)$ yields the compactness of the operator $\mathbf{T}$.

As a direct consequence of those theorems recalled above, we conclude the following classic result.

Theorem 9.13 *There exist a sequence of positive numbers* $\{\lambda_j\}$*,* $\lambda_j \to \infty$*, and a sequence of functions*

$$\{u_j\} \subset H^1_0(\Omega) \cap C^\infty(\Omega)$$

that make up an orthonormal basis of $L^2(\Omega)$*, such that*

$$-\Delta u_j = \lambda_j u_j \text{ in } \Omega. \tag{9.29}$$

Proof We have already checked that the operator $\mathbf{T}$ is compact and self-adjoint. It is, in addition, positive which means that

$$\langle \mathbf{T}u, u\rangle_{L^2(\Omega)} = \|\mathbf{T}u\|^2_{H^1(\Omega)} \ge 0$$

for every $u \in L^2(\Omega)$. By Theorems 6.4 and 6.3, we can conclude the existence of a sequence of non-vanishing eigenvalues $\{1/\lambda_j\}$, which are positive because so is $\mathbf{T}$, converging to zero, and a sequence of corresponding eigenfunctions $\{u_j\}$, i.e. (9.29) holds. Finally, from the regularity results from the previous section, which may be restricted to arbitrary compact, smooth subdomains of Ω, utilized in a recursive way through (9.29), we conclude the (interior) smoothness of each eigenfunction u_j. □

For obvious reasons, the numbers λ_j in (9.29) are called eigenvalues of the Laplace operator (under vanishing Dirichlet boundary conditions) in Ω, while the corresponding u_j are the associated eigenfunctions.

It is interesting to realize, as we described in the Introduction to this chapter, that solutions for the problem

$$-\Delta u = \lambda u \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega,$$

correspond to critical functions for the functional

$$\int_\Omega \left(\frac{1}{2}|\nabla u|^2 - \lambda\frac{1}{2}u^2\right)\, d\boldsymbol{x}.$$

We now know that when $\lambda \le 0$, the corresponding integrand

$$(u, \mathbf{u}) \mapsto \frac{1}{2}|\mathbf{u}|^2 - \lambda \frac{1}{2} u^2$$

is convex, or even strictly convex if $\lambda < 0$, and hence the unique critical function would be the unique minimizer which is the trivial function. For $\lambda > 0$, this is no longer true. However, the associated variational principle might not be well-posed in the sense that the infimum might decrease to $-\infty$. To recover a meaningful problem, one needs to limit the size of the competing functions u

$$\text{Minimize in } u \in H_0^1(\Omega): \quad \frac{1}{2} \int_\Omega |\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x}$$

subject to

$$\int_\Omega u(\boldsymbol{x})^2 \, d\boldsymbol{x} = 1.$$

For such a variational problem, the set $\mathcal{A}$ of admissible functions would be

$$\mathcal{A} = \{u \in H_0^1(\Omega) : \|u\|_{L^2(\Omega)} = 1\}.$$

Because of the compact injection of $H_0^1(\Omega)$ into $L^2(\Omega)$, the set $\mathcal{A}$ is weakly closed, and hence there is a unique global minimizer $u_1 \in H_0^1(\Omega)$ for such a problem. Recall our discussion in Sect. 9.2. Indeed, Theorem 9.13 ensures that there is a whole sequence of such values.

9.9 Duality for Sobolev Spaces

Even though duality is quite well-known and useful for Lebesgue spaces, it not so for Sobolev spaces. The main reason is that Sobolev spaces are the fundamental function spaces for variational problems and PDEs involving weak derivatives, and in this context duality does not play such a prominent role.

Possibly, the most used dual for a Sobolev space is the dual space of $H_0^1(\Omega)$, which is designated by $H^{-1}(\Omega)$ to stress that functions in this space have one "negative" weak derivative in the sense that they can be identified with derivatives of functions in $L^2(\Omega)$ (which do not admit, in general, weak derivatives).

Proposition 9.5 *Let $T \in H^{-1}(\Omega)$. Then there is $\mathbf{F} \in L^2(\Omega; \mathbb{R}^N)$ such that, formally, $T = \operatorname{div} \mathbf{F}$, that is to say*

$$\langle T, u\rangle = \int_\Omega \mathbf{F}(\boldsymbol{x}) \cdot \nabla u(\boldsymbol{x})\, d\boldsymbol{x} \tag{9.30}$$

for all $u \in H_0^1(\Omega)$.

As a matter of fact, the identification of the dual of a Hilbert space with itself leads to a more explicit identification: for $T \in H^{-1}(\Omega)$, there is a unique $v \in H_0^1(\Omega)$ such that

$$\langle T, u\rangle = \int_\Omega \nabla v(\boldsymbol{x}) \cdot \nabla u(\boldsymbol{x})\, d\mathbf{x}, \tag{9.31}$$

so that $\mathbf{F} = \nabla v$ for a unique $v \in H_0^1(\Omega)$. Moreover

$$\|T\|_{H^{-1}(\Omega)} = \|v\|_{H^1(\Omega)}.$$

Formally, (9.31) can be written as

$$-\Delta v = T, \quad \|T\|_{H^{-1}(\Omega)} = \|v\|_{H^1(\Omega)}.$$

On the other hand, every field $\mathbf{F} \in L^2(\Omega; \mathbb{R}^N)$ defines an element T of $H^{-1}(\Omega)$ through the integral (9.30). Again, it can be represented by a unique $v \in H_0^1(\Omega)$ such that

$$\int_\Omega \mathbf{F}(\boldsymbol{x}) \cdot \nabla u(\boldsymbol{x})\, d\boldsymbol{x} = \int_\Omega \nabla v(\boldsymbol{x}) \cdot \nabla u(\boldsymbol{x})\, d\boldsymbol{x} \tag{9.32}$$

for all $u \in H_0^1(\Omega)$. Condition (9.32) amounts to

$$\operatorname{div}(\mathbf{F} - \nabla v) = 0 \text{ in } \Omega, \quad v \in H_0^1(\Omega),$$

i.e. the function $v \in H_0^1(\Omega)$ representing the element $\operatorname{div} \mathbf{F} \in H^{-1}(\Omega)$ is the unique minimizer v of the quadratic problem

$$\text{Minimize in } u \in H_0^1(\Omega) : \quad \frac{1}{2}\int_\Omega |\mathbf{F}(\boldsymbol{x}) - \nabla u(\boldsymbol{x})|^2\, d\boldsymbol{x}.$$

Even more generally, for a given $T \in H^{-1}(\Omega)$, it is identified as the unique solution of the quadratic problem (recall Lemma 2.7)

$$\text{Minimize in } u \in H_0^1(\Omega) : \quad \frac{1}{2}\|u\|^2 - \langle T, u\rangle.$$

The optimality condition for the minimizer v of this problem becomes exactly (9.31). In practice, elements T of the dual space $H^{-1}(\Omega)$ are manipulated via its unique representative $v \in H_0^1(\Omega)$ given by these conditions.

There are analogous identifications for elements of dual spaces $W^{-1,q}(\Omega)$ for $W_0^{1,p}(\Omega)$ and conjugate exponent q, although the identification cannot go as far as with the Hilbert space $H_0^1(\Omega)$.

9.10 Exercises

1. Let $u(\boldsymbol{x})$ be a smooth function defined on a domain $\Omega \subset \mathbb{R}^N$ limited by a smooth (C^2-) manifold $\partial\Omega$ such that

$$u|_{\partial\Omega} \equiv 0.$$

Let smooth functions $(\boldsymbol{X}(\boldsymbol{x}), X_N(\boldsymbol{x}))$ be defined in a neighborhood $\mathbf{U}$ of any point in $\partial\Omega$ in such a way that at every point of $\mathbf{U}$, gradients $\nabla X_i(\boldsymbol{x})$ for $i = 1, 2, \ldots, N-1$ are tangent to $\partial\Omega$ while $\nabla X_N(\boldsymbol{x})$ is the unit, outer normal $\mathbf{n}(\boldsymbol{x})$ to $\partial\Omega$.Put

$$u(\boldsymbol{x}) = U(\boldsymbol{X}(\boldsymbol{x}), X_N(\boldsymbol{x})).$$

(a) Argue that

$$\nabla^2 u = \frac{\partial^2 U}{\partial X_N^2}(\boldsymbol{X}, X_N)\nabla X_N \otimes \nabla X_N + \frac{\partial U}{\partial X_N}(\boldsymbol{X}, X_N)\nabla^2 X_N$$

over $\mathbf{U}$.

(b) Proceed by density to prove (9.24).

2. Show that the best constant in the Poincaré's inequality in $H_0^1(\Omega)$ is the inverse of the first positive eigenvalue of the operator $-\Delta$ by examining the constrained variational problem

$$\text{Minimize in } u(\boldsymbol{x}) \in H_0^1(\Omega): \quad \int_\Omega |\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x}$$

under the integral constraint

$$\int_\Omega u(\boldsymbol{x})^2 \, d\boldsymbol{x} = \text{positive constant}.$$

3. Examine the various results in Sect. 9.6 and how they should be adapted for dimensions $N = 1$ and $N = 2$.

4. Prove that if for a C^1-function $f(x)$, we know that

$$0 < pf(x) \le xf'(x), \quad |x| \ge R > 0, p > 0,$$

then there are constants $a > 0$ and b such that

$$f(x) \ge a|x|^p + b$$

for all real x.

5. Explore the eigenvalues and eigenfunctions of the Laplace operator under Neumann boundary conditions following the same steps as with the case of Dirichlet boundary conditions. The following fact (which we have not proved in the text)

$$\|u\|_{L^2(\Omega)} \le C\|\nabla u\|_{L^2(\Omega;\mathbb{R}^N)}$$

is valid for functions in the space $H^1(\Omega) \cap L^2_0(\Omega)$ where

$$L^2_0(\Omega) = \{u \in L^2(\Omega) : \int_\Omega u(\mathbf{x})\, d\mathbf{x} = 0\}.$$

This inequality is known as Wirtinger's inequality.

6. (a) Explore the variational constrained problem

$$\text{Minimize in } u \in H^1(\Omega) : \quad \int_\Omega |u(\mathbf{x})|^2\, d\mathbf{x}$$

subject to

$$\int_\Omega |\nabla u(\mathbf{x})|^2\, d\mathbf{x} = 1.$$

Is there an optimal solution? Look for an easy counterexample in dimension 1.

(b) Consider the variational problem

$$\text{Maximize in } u \in H^1_0(\Omega) : \quad \int_\Omega |u(\mathbf{x})|^2\, d\mathbf{x}$$

subject to

$$\int_\Omega |\nabla u(\mathbf{x})|^2\, d\mathbf{x} \le 1.$$

Is there an optimal solution?

7. For a real function

$$f(t) : [0, +\infty) \to [0, \infty), \quad (\text{ like } f(t) = \sqrt{t}),$$

use Theorem 9.13 to define $f(\Delta)$ through the spectrum of Δ.

8. Complete the induction argument of Lemma 9.1.
9. Keep track of the various constants in the recursive process associated with Theorem 9.2, and conclude that it is independent of the domain Ω.
10. Show how in the critical case $p = N$, the recursive process associated with the proof of Theorem 9.2 can be pushed to infinity.
11. Suppose u is a smooth minimizer for the Dirichlet integral

$$\frac{1}{2} \int_{\Omega} |\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x}, \quad \Omega \subset \mathbb{R}^N,$$

under some (non-null) Dirichlet boundary conditions. Use inner variations of the form

$$U_\epsilon(\boldsymbol{x}) = u(\boldsymbol{x} + \epsilon \Phi(\boldsymbol{x})), \quad \Phi(\boldsymbol{x}) : \Omega \to \mathbb{R}^N, \ \Phi|_{\partial\Omega} = \mathbf{0},$$

to deduce the corresponding optimality conditions.

12. The basic strategy of the direct method can be used in more general situations. Consider the following one. It is a matter of

$$\text{Minimize in } \mathbf{F} : \quad E(\mathbf{F}) = \int_{\Omega} \left[\frac{1}{2} \mathbf{F}(\boldsymbol{x})^T \mathbb{A}(\boldsymbol{x}) \mathbf{F}(\boldsymbol{x}) + \mathbf{A}(\boldsymbol{x}) \cdot \mathbf{F}(\boldsymbol{x}) \right] d\boldsymbol{x}$$

where $\mathbf{F}(\boldsymbol{x}) : \Omega \subset \mathbb{R}^N \to \mathbb{R}^N$, and

$$\operatorname{div} \mathbf{F} = 0 \text{ in } \Omega, \quad \mathbf{F} \cdot \boldsymbol{n} = 0 \text{ on } \partial\Omega.$$

(a) Define in a suitable way the space where a weak null-divergence can be enforced, together with the normal component on the boundary.
(b) Apply the direct method to the proposed problem.
(c) Derive the form of optimality conditions in this case.

Appendix A
Hints and Solutions to Exercises

A.1 Chapter 1

1. Computations are not hard if performed with some care.
2. No comment.
3. (a) If there is no constraint to be respected, the minimizer will be achieved by minimizing the integrand pointwise for each $\boldsymbol{x} \in \Omega$: the optimal function will be

$$\overline{u}(\boldsymbol{x}) = \min_{u} F(\boldsymbol{x}, u),$$

and optimality conditions will be given by

$$\frac{\partial F}{\partial u}(\boldsymbol{x}, u) = 0 \text{ for } \boldsymbol{x} \in \Omega.$$

As soon as there are global conditions to be preserved, this process can hardly yield the optimal solution. Even a condition like

$$\int_{\Omega} u(\boldsymbol{x})\, d\boldsymbol{x} = u_0$$

spoils the above procedure, as it is shown next.

(b) A bit of reflection leads to the minimizer $\chi_{\mathbf{A}}$ where the set $\mathbf{A}$ is described by the condition on constant M so that

$$\mathbf{A} = \{\boldsymbol{x} \in \Omega : f(\boldsymbol{x}) \le M\}, \quad |\mathbf{A}| = s|\Omega|.$$

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4

4. If

$$h(x) : [0, L] \to \mathbb{R}^+, \quad h(0) = h(L) = 0$$

is an arbitrary path joining the departure point (0, 0) and the arrival point $(L, 0)$, its cost will be given by the integral

$$\int_0^L e^{-ah(x)} \sqrt{1 + h'(x)^2}\, dx$$

where

$$\sqrt{1 + h'(x)^2}\, dx$$

is the element of arc-length.

5. We have already used earlier in the chapter that

$$T = \int dt = \int \frac{ds}{v}.$$

In this particular situation we would have

$$T = \int_0^L \boldsymbol{n}(x, y(x)) \sqrt{1 + y'(x)^2}\, dx$$

under the end-point conditions

$$y(0) = y_0, \quad y(L) = y_L.$$

6. It is just a matter of bearing in mind that

$$y'(x)x'(y) = 1, \quad dx = x'(y)\, dy$$

to find that the resistance functional can be written in the form

$$\int_H^h \frac{y}{1 + x'(y)^2}\, dy.$$

7. From elementary courses of Calculus, it is known that the surface of revolution generated by the graph of the a function $u(x)$ between $x = 0$ and $x = L$ is given by the integral

$$S = 2\pi \int_0^L u(x) \sqrt{1 + u'(x)^2}\, dx.$$

8. A curve $\sigma(t)$ with image contained in the cylinder of equation

$$x^2 + y^2 = 1$$

can obviously parametrized in the form

$$\sigma(t) = (\cos t, \sin t, z(t))$$

with $t \in [0, \theta]$, with $\theta \leq \pi$, if

$$P \mapsto t = 0, z(0) = 0, \quad Q \mapsto t = \theta, z(\theta) = L \in \mathbb{R},$$

if the points P and Q are, respectively, the initial and final points and L can be positive or negative. The functional providing the length of such a curve is then

$$\int_0^\theta \sqrt{1 + z'(t)^2}\, dt.$$

For the sphere, one can use spherical coordinates and write

$$\sigma(t) = (\sin z(t) \cos t, \sin z(t) \sin t, \cos z(t)),$$
$$z(0) = \pi/2, z(\theta) = \phi \in [0, \pi],$$

if, without loss of generality,

$$(0, \pi/2, (\theta, \phi), \quad 0 \leq \theta \leq \pi,$$

are the spherical coordinates of the initial and final points.

9. It is easy to write the functional

$$\int_\Omega \sqrt{1 + u_x(x, y)^2}\sqrt{1 + u_y(x, y)^2}\, dx\, dy.$$

Clearly a function for which $u_x = u_y = 0$, i.e. a constant function, is a minimizer of this functional. It would be hard to give another one.

10. It is clear that all possible parametrizations of the path $\mathbf{u}(t)$ are of the form

$$\mathbf{v}(s) = \mathbf{u}(\gamma(s)), \quad t = \gamma(s),$$

with

$$\gamma(s) : [0, 1] \to [0, 1], \quad \gamma(0) = 0, \gamma(1) = 1.$$

The variational problem would be

$$\text{Minimize in } \gamma: \quad \int_0^1 F(\mathbf{u}(\gamma(s)), \gamma'(s)\mathbf{u}'(\gamma(s)))\, ds.$$

If γ is assumed bijective and smooth, then a change of variables leads to the problem of minimizing in γ

$$\int_0^1 F\left(\mathbf{u}(t), \frac{1}{\gamma'(t)}\mathbf{u}'(t)\right)\gamma'(t)\, dt.$$

Here we have replaced γ by its inverse γ^{-1}. We therefore see that those integrands $F(\mathbf{u}, \mathbf{v})$ that are homogeneous of degree 1 in $\mathbf{v}$ are parametrization-independent. A fundamental example is

$$F(\mathbf{u}, \mathbf{v}) = |\mathbf{v}|$$

whose associated functional is the length of the curve.

A.2 Chapter 2

1. That these two sets are vector spaces is a direct consequence of the triangular inequality. As in other initial examples in the chapter, $L^\infty(\Omega)$ is seen to be complete. The case of $L^1(\Omega)$ is included in the proof of Proposition 2.2. It is a good idea to review that proof for the case $p = 1$.
2. This is clear if we notice that every continuous function g in $\overline{\Omega}$ always belong to $L^q(\Omega)$, even if $q = \infty$. More directly, simply use Hölder's inequality for $q < \infty$.
3. This kind of examples is easy to design once the underlying difficulty is understood. For instance, take

$$u_j(x) = \int_0^x j\chi_{(0,1/j)}(y)\, dy, \quad u_j'(x) = j\chi_{(0,1/j)}(x), u_j(0) = 0,$$

where $\chi_{(0,1/j)}$ is the characteristic function of the interval $(0, 1/j)$ in the unit interval $(0, 1)$. It is easy to check that $\{u_j\}$ is uniformly bounded in $W^{1,1}(0, 1)$ but it is not equicontinuous because the limit of any subsequence cannot be continuous.
4. (a) This is elementary.
   (b) By using the double-angle formula, it is not difficult to conclude that

$$\int_a^b u_j^2(x)\, dx \to \frac{b-a}{2}.$$

In terms of weak convergence, we see that

$$u_j \rightharpoonup 0, \quad u_j^2 \rightharpoonup \frac{1}{2},$$

which is a clear indication of the unexpected behavior of weak convergence with respect to non-linearities.

5. Under the hypotheses assumed on $\{f_j\}$, we know that there is some subsequence j_k and $\overline{f} \in L^p(J)$ with

$$f_{j_k} \rightharpoonup \overline{f} \text{ in } L^p(J).$$

But the convergences in arbitrary subintervals imply that

$$\int_a^b f(x)\,dx = \int_a^b \overline{f}(x)\,dx.$$

Since a and b are arbitrary, we can conclude that $f \equiv \overline{f}$ as elements in $L^p(J)$. Since the convergence in subintervals takes place for the full sequence (no subsequences), the same is correct at the level of weak convergence.

6. Let $\{\mathbf{x}'_j\}$ be a Cauchy sequence in $\mathbb{E}'$. Proceed in two steps:

   (a) Show that the limit $\{\mathbf{x}'_j(\mathbf{x})\}$ exists for every individual $\mathbf{x} \in \mathbb{E}$. Define $\mathbf{x}'(\mathbf{x})$, as the limit. $\mathbf{x}'$ is linear by definition.
   (b) Show that the set of numbers $\{\|\mathbf{x}'_j\|\}$ is bounded, and conclude that $\mathbf{x}' \in \mathbb{E}'$.

7. The proof of Proposition 2.7 can be finished after Exercise 5 above.
8. The vector situation in Proposition 2.8 is similar to the scalar case. It is a good way to mature these ideas.
9. If one admits that

$$\mathbf{x} = \sum_{j=1}^{\infty} a_j \mathbf{x}_j \in \mathbb{H}$$

because

$$\|\mathbf{x}\|^2 \leq \sum_j a_j^2 < \infty,$$

then the other conclusions are a consequence of the first part of the proposition. Based on the summability of the series of the a_j's, it is easy to check that the

sequence of partial sums

$$\{\sum_{j=1}^{N} a_j \mathbf{x}_j\}_N$$

is convergent.

10. (a) The same example in Exercise 3 is valid to show that $\mathbb{E}_1$ is not complete.
    (b) It is not difficult to realize that $\|\mathbf{T}_g\| = \|g\|_\infty$. In fact, if we regard $\mathbb{E}_1$ as a (dense) subspace of $L^1(0, 1)$, then it is clear that $\|\mathbf{T}_g\| = \|g\|_\infty$.
    (c) It suffices to take $f = g/|g|$ on the subset of $(0, 1)$ where g does not vanish.
    (d) It is obvious that $\delta_{1/2}$ is a linear, continuous functional on $\mathbb{E}_\infty$. However, it does not belong to $\mathbb{E}_1'$ because functions can take on arbitrarily large values off $\{1/2\}$ without changing its value at this point.
    (e) The subspace $\mathbb{H}$ is the zero-set of the linear, continuous functional $\mathbf{T}_g$ for $g \equiv 1$. Since by the previous items, $\mathbf{T}_1$ belongs both to $\mathbb{E}_1'$ and $\mathbb{E}_\infty'$, we have our conclusion.
11. (a) This is a direct consequence of the triangular inequality and works as in any vector space; namely

$$\|\mathbf{x}\| \le \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\|, \quad \|\mathbf{y}\| \le \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{x}\|,$$

    for any two pair of vectors $\mathbf{x}$, $\mathbf{y}$, and so

$$|\|\mathbf{x}\| - \|\mathbf{y}\|| \le \|\mathbf{x} - \mathbf{y}\|.$$

    (b) Let

$$\mathbf{z} = \sum_i z_i \mathbf{e}_i, \quad \{\mathbf{e}_i : i\} \subset \mathbb{R}^n, \text{ the canonical basis.}$$

    Then, by linearity,

$$\mathbf{T}\mathbf{z} = \sum_i z_i \mathbf{T}\mathbf{e}_i,$$

    and

$$\|\mathbf{T}\mathbf{z}\| \le \sum_i |z_i| \|\mathbf{T}\mathbf{e}_i\| \le \|\mathbf{z}\|_\infty \sum_i \|\mathbf{T}\mathbf{e}_i\|.$$

    (c) This is immediate: for every k, j larger than some j_0, we will have

$$\|\mathbf{x}_j - \mathbf{x}_k\| \le 1$$

if $\{\mathbf{x}_j\}$ is a Cauchy sequence. In particular,

$$\|\mathbf{x}_j\| \leq \|\mathbf{x}_j - \mathbf{x}_{j_0}\| + \|\mathbf{x}_{j_0}\| \leq 1 + \|\mathbf{x}_{j_0}\|,$$

for every j larger than j_0.

(d) It is easy to realize, if $\{\mathbf{e}_i\}$ is the canonical basis of $\mathbb{R}^n$, that if

$$\mathbf{x} = \sum_i x_i \mathbf{e}_i,$$

then

$$\|\mathbf{x}\| \leq \sum_i |x_i| \|\mathbf{e}_i\| \leq \sum_i \|\mathbf{e}_i\| \|\mathbf{x}\|_\infty = M \|\mathbf{x}\|_\infty.$$

12. (a) It is clear that if $\mathbf{x} \in \ell^p$ for some positive p, i.e. the series

$$\sum_j |x_j|^p < \infty, \quad \mathbf{x} = (x_1, x_2, \dots),$$

then all the terms, except a finite number, must be less than 1, and hence

$$|x_j|^q < |x_j|^p, \quad \text{and } \mathbf{x} \in \ell^q.$$

Moreover, there is always a finite index set $\mathbf{K} \subset \mathbb{N}$ such that

$$\|\mathbf{x}\|_\infty = |x_k|, \quad |x_j| \leq |x_k|,$$

for $k \in \mathbf{K}$, and all j. Hence

$$\frac{\|\mathbf{x}\|_q}{\|\mathbf{x}\|_\infty} = \left(\sum_j \frac{|x_j|^q}{|x_k|^q} \right)^{1/q}$$

share the limit with

$$\#(\mathbf{K})^{1/q} \to 1, \quad \text{as } q \to \infty.$$

(b) This is similar to the previous item.

13. (a) It is clear that

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1.$$

However, the elements

$$\mathbf{x}_k \equiv \sum_{i\in I_k} \mathbf{e}_i$$

for a sequence of finite sets I_k with cardinal $k \to \infty$, are such that

$$\|\mathbf{x}_k\|_1 = k, \quad \|\mathbf{x}_k\|_\infty = 1.$$

(b) The index set I is infinite (because $\mathbb{E}$ has infinite dimension), and select a copy of $\mathbb{N}$ within I. Define a linear functional $\mathbf{T}$ over $\mathbb{E}$ by defining it on the basis $\{\mathbf{e}_i\}$ and putting

$$\mathbf{T}(\mathbf{e}_n) = n\|\mathbf{e}_n\|, n \in \mathbb{N}, \quad \mathbf{T}(\mathbf{e}_i) = 0, i \in I \setminus \mathbb{N}.$$

This linear functional cannot be continuous.

14. The given function $\|\cdot\|$ is a norm, because the only solution of the problem

$$u(t) + u'(t) = 0 \text{ in } [0, 1], \quad u(0) = 0,$$

is the trivial function. The other conditions are inherited from the similar properties of the absolute value. In particular,

$$\|u\| \le \|u\|_\infty + \|u'\|_\infty.$$

For the reverse inequality, bear in mind that

$$|u(t)| \le \|u\| + \int_0^t |u(s)|\,ds,$$

$$|u'(t)| \le \|u\| + \int_0^t |u'(s)|\,ds,$$

and use Gronwall's inequality.

15. (a) implies (b) is elementary. For the reverse implication, use a sequence $\{\rho_j\}$ in $L^1(\mathbb{R})$, as in Sect. 2.29 with the property

$$\mathbf{T}(\rho_j) * v = \mathbf{T}(\rho_j * v) \to \mathbf{T}(v)$$

for all v. $\{\rho_j\}$ is called an approximation of the identity in $L^1(\mathbb{R})$. On the other hand, show that some subsequence of $\{\mathbf{T}(\rho_j)\}$ converges weakly in $L^p(\mathbb{R})$ to some w.

16. Note that for fixed $\mathbf{x} \in \mathbb{H}$, the convergence of the series

$$\sum_j \langle \mathbf{x}_j, \mathbf{x} \rangle \mathbf{x}_j$$

implies that

$$\langle \mathbf{x}_j, \mathbf{x} \rangle \to 0,$$

for all $\mathbf{x} \in \mathbb{H}$.

17. All conditions of an inner product are fine, except for the last one. Note that if a non-null function $f(\boldsymbol{x})$ is orthogonal (in $L^2(\Omega)$) to the full set $\{w_i\}$, then the inner-product with a double integral would vanish.

18. Integrate by parts in the product

$$\int_{-1}^{1} L_j(t) L_k(t)\, dt, \quad j > k,$$

as many times as necessary, charging derivatives on L_k and checking that contributions at the end-points vanish, until you see that the integral vanishes.

(a) They are asking about the projection of t^3 onto the subspace generated by $\{L_0, L_1, L_2\}$, namely

$$\pi t^3 = \sum_{i=0}^{2} \frac{\langle t^3, L_i \rangle}{\langle L_i, L_i \rangle} L_i(t).$$

More specifically, the minimum sought is

$$\int_{-1}^{1} |t^3 - \pi t^3|^2\, dt.$$

(b) The maximum is attained for

$$p = \frac{1}{\|t^3 - \pi t^3\|} (t^3 - \pi t^3).$$

19. (a) It suffices to realize that the leading coefficient of all the P_j's is unity, as a consequence of the Gram-Schmidt orthogonalization process, and so $t P_{j-1}(t) - P_j(t)$ has, at most, degree $j - 1$.

(b) One must show that

$$\langle P_j(t) - t P_{j-1}(t), P_k(t) \rangle = 0, \quad k \le j - 3.$$

Indeed,

$$\langle tP_{j-1}(t), P_k(t)\rangle = \langle P_{j-1}(t), tP_k(t)\rangle = 0$$

if $k \leq j-3$. This implies that

$$P_j(t) - tP_{j-1}(t) = a_j P_{j-1}(t) + b_j P_{j-2}(t).$$

The inner product with P_{j-2}, after the use of the previous item, yields the sign of b_j.

(c) Argue first that P_j ought to have at least one real root in J because $\langle P_j, 1\rangle = 0$. If we let

$$t_1 < t_2 < \cdots < t_q$$

be the roots of P_j in J in which P_j changes sign and set

$$P(t) = \Pi_{i=1}^{q}(t - t_i)$$

argue that $\langle P_j, P\rangle$ cannot vanish, which is absurd if $q < j$.

20. It is a simple exercise to check that the complex system of exponential is orthonormal. To see that it is a basis for $\mathbb{H}$, separate in real and imaginary parts.
21. Proceed in various steps:

(a) Show first that $\{\psi_{m,n}\}$ is orthonormal by considering the two cases (m_1, n), (m_2, n), first; and then (m_1, n_1), (m_2, n_2).

(b) Argue that every function in $L^2(\mathbb{R})$ can be represented as a linear combination of the Haar system of wavelets in the following way.

(i) Define the subspace $\mathbb{H}_j$ as the set of square-integrable functions which are constant on dyadic intervals of length 2^{-j}. Note that

$$\cdots \subset \mathbb{H}_{-2} \subset \mathbb{H}_{-1} \subset \mathbb{H}_0 \subset \mathbb{H}_1 \subset \mathbb{H}_2 \subset \dots,$$

and that

$$\cap_j \mathbb{H}_j = \{0\}.$$

(ii) Show that $\mathbb{H} = \cup_j \mathbb{H}_j$ is dense in $L^2(\mathbb{R})$. Characteristic functions of the form $f(x) = \chi_{[a,b]}(x)$ can be approximated by elements of $\mathbb{H}$ if we put

$$a = \frac{k}{2^n} - a_1, \quad b = \frac{l}{2^n} + b_1, \quad a_1, b_1 < 2^{-n},$$

$$g(x) = \chi_{[k/2^n, l/2^n]}(x) \in \mathbb{H}, \quad \|f - g\| \leq 2^{1-n}.$$

Since $\mathbb{H}$ is a subspace, linear combination of characteristics (simple or step functions) belong to $\mathbb{H}$ too. Since simple functions are dense in $L^2(\mathbb{R})$, so is $\mathbb{H}$.

22. Though a complete and clear answer to this problem would require further knowledge about wavelet analysis, take the previous exercise as a model, and write conditions on ψ in terms of the density in $L^2(\mathbb{R})$ of a certain subspace associated with ψ.
23. Such complement is spanned by functions $u(t)$ such that

$$\int_J (u(t)v(t) + u'(t)v'(t))\,dt = 0, \quad v \in H_0^1(J).$$

By integrating by parts on the second term, we deduce that $u = u''$.

24. The same procedure shown in (2.27) is valid in more generality for a linear, self-adjoint, differential operator $\mathbb{L}$. Indeed, if

$$\mathbb{L}u_k + \omega_k^2 u_k = 0 \text{ in } (0, 1), \quad u_k(0) = u_k(1) = 0,$$

for a certain sequence of (pair-wise different) numbers ω_k so that u_k is not the trivial function, then

$$\begin{aligned}\omega_k^2 \int_0^1 u_k(x)u_j(x)\,dx &= -\int_0^1 \mathbb{L}u_k(x)u_j(x)\,dx \\ &= -\int_0^1 u_k(x)\mathbb{L}u_j(x)\,dx \\ &= \omega_j^2 \int_0^1 u_k(x)u_j(x)\,dx.\end{aligned}$$

If $\omega_k \neq \omega_j$, we conclude that u_k and u_j must be orthogonal. The self-adjointness of $\mathbb{L}$ (see Chap. 5 below) means, through integration by parts, precisely that

$$\int_0^1 \mathbb{L}u(x)v(x)\,dx = \int_0^1 u(x)\mathbb{L}v(x)\,dx,$$

under the end-point conditions

$$u_k(0) = u_k(1) = 0.$$

Operators of the form

$$\mathbb{L}u(x) = -(p(x)u'(x))' + q(x)u(x)$$

for arbitrary functions $p(x) > 0$ and $q(x)$ are valid for these computations to hold (Sturm-Liouville eigenvalue problems).

25. For the triangle inequality, note that the function $\phi(r) = r/(1 + r)$ is strictly increasing for positive r. Then

$$\begin{aligned}\phi(|f - g|) \leq& \phi(|f - h| + |h - g|)\\ =& \frac{|f - h| + |h - g|}{1 + |f - h| + |h - g|}\\ \leq& \phi(|f - h|) + \phi(|h - g|).\end{aligned}$$

For the completeness, reduce the case to $L^1(J)$.

26. (a) This is easy to argue because for positive numbers a and b, we always have

$$\sqrt{a + b} \leq \sqrt{a} + \sqrt{b}.$$

(b) However $\|\cdot\|_{1/2}$ is not a norm because it does not respect the triangular inequality. To check this, consider the situation for χ, the characteristic function of the interval $(0, 1/2)$ in $(0, 1)$, and $2 - \chi$.

(c) As in the first item above, it is easy to check that $d(f, g)$ is indeed a distance. The completeness reduces to the same issue in $L^p(0, 1)$ for $p \geq 1$.

27. Define

$$\begin{aligned}\langle \mathbf{x}, \mathbf{y}\rangle =& \frac{1}{4}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2)\\ =& \frac{1}{2}\|\mathbf{x}\|^2 + \frac{1}{2}\|\mathbf{y}\|^2 - \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|^2,\end{aligned}$$

and prove that it is an inner-product whose associated norm is $\|\cdot\|$, through the following steps:

(a) Check through the parallelogram identity that

$$2\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x} + \mathbf{z}, \mathbf{y}\rangle + \langle \mathbf{x} - \mathbf{z}, \mathbf{y}\rangle.$$

(b) By writing

$$\langle \mathbf{x} + \mathbf{z}, \mathbf{y}\rangle = \frac{1}{2}\left(\langle \mathbf{x} + \mathbf{z} + \mathbf{w}, \mathbf{y}\rangle + \langle \mathbf{x} + \mathbf{z} - \mathbf{w}, \mathbf{y}\rangle\right),$$

and choosing $\mathbf{w}$ in an appropriate way, conclude that

$$\langle \mathbf{x} + \mathbf{z}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{z}, \mathbf{y}\rangle.$$

(c) Show that

$$\langle 2\mathbf{x}, \mathbf{y}\rangle = 2\langle \mathbf{x}, \mathbf{y}\rangle,$$

and then, writing an arbitrary λ in base 2, through the previous item, conclude that

$$\langle \lambda\mathbf{x}, \mathbf{y}\rangle = \lambda\langle \mathbf{x}, \mathbf{y}\rangle.$$

28. If we rely on the similar result for the one-dimensional situation, one can deduce the same result for the closure of the subspace of $L^2([-\pi, \pi]^N; \mathbb{C})$ made up of linear combinations of the form

$$\sum_l \Pi_k f_{kl}(x_k), \quad f_{kl} \in L^2([-\pi, \pi]; \mathbb{C}).$$

It is a standard separation-of-variables process.

29. The projection theorem Proposition 2.11 can be applied directly to this situation under the following elements:

(a) $\mathbb{H}$ is $H_0^1(0, 1)$ while we take $\mathbf{x}$, the trivial function. The norm in $\mathbb{H}$ is precisely

$$\|u\| = \int_0^1 u'(x)^2\, dx.$$

(b) $\mathbb{K}$ is the subset of $\mathbb{H}$ given by

$$\mathbb{K} = \{u \in H_0^1(0, 1) : u(x) \le \phi(x)\}$$

which is clearly convex.

One can then conclude the existence of a unique function $\overline{u} \in \mathbb{K}$ realizing the minimum of the problem. This function is characterized by the condition

$$\int_0^1 \overline{u}(x)[u(x) - \overline{u}(x)]\, dx \ge 0$$

for every $u \in \mathbb{K}$.

30. The upper bounds given in the statement ensure that the functional E is well-defined in $H_0^1(0, 1)$. According to Definition 2.13, we need to calculate the derivative of the section

$$t \mapsto E(u+tv) = \int_0^1 [\psi(u'(t)+tv'(x))+\phi(u(x)+tv(x))]\, dx, \quad v \in H_0^1(0, 1),$$

which is explicitly given by the formula

$$\langle E'(u), v\rangle = \int_0^1 [\psi'(u'(x))v'(x) + \phi'(u(x))v(x)]\, dx.$$

Lemma 2.7 informs us that the derivative $E'(u)$ will be the unique minimizer of the problem

$$v \in H_0^1(0,1) \mapsto \int_0^1 \left[\frac{1}{2}v'(x)^2 - \psi'(u'(x))v'(x) - \phi'(u(x))v(x)\right] dx.$$

31. This is a typical situation for the application of Theorem 2.1 of Chap. 2, in a similar way as in Sect. 2.10 of the same chapter.

A.3 Chapter 3

1. This is immediate from the statement of Corollary 3.2.
2. If we let a, b, and c be those three numbers, respectively, then it is immediate to check that $a \le b \le c$. For this last inequality, note that there is always $\mathbf{x}' \in \mathbb{E}'$ with

$$\left\langle \mathbf{x}', \mathbf{x}_j \right\rangle = \lambda_j \|\mathbf{x}_j\|$$

for every particular choice of numbers $\lambda_j \in [-1, 1]$. In particular, there is always one such $\hat{\mathbf{x}}'$ with

$$\langle \hat{\mathbf{x}}', \sum_j \lambda_j \mathbf{x}_j\rangle = \|\sum_j \lambda_j \mathbf{x}_j\|$$

if λ_j, with $|\lambda_j| = 1$, are the numbers furnishing the sup for a.
3. This is elementary if argued through sequences.
4. This is a fine exercise in Mutivariate Calculus.
5. (a) This is an interesting exercise of the manipulation of the convexity inequality to end with the sought inequality.
   (b) Check that the previous proof in a finite-dimensional scenario is valid in an infinite-dimensional situation too. The application to Proposition 3.4 is immediate.
6. Multiply (3.5) by an arbitrary $v \in H_0^1(J)$, and perform an integration by parts.
7. (a) This is standard. Check first that $\langle u, v\rangle$ is a seminorm, and then that the associated norm only vanishes at the trivial function.

(b) The corresponding Lax-Milgram lemma is concerned with the minimization problem for the functional

$$\frac{1}{2}\int_0^1 tu'(t)^2\,dt$$

over $\mathbb{H}$. It is enlightening to check that there cannot be a minimizer, since if there would, it would be a solution of the Euler-Lagrange problem

$$-[tu'(t)]' = 0 \text{ in } (0,1), \quad u(0) = u(1) = 0.$$

It is elementary to check that there is no solution for this problem. The whole point is that the given bilinear form is not coercive as the coercivity constant t degenerates as t moves closer to the origen.

8. The linearity of such a mapping is clear. For the continuity, use $\mathbf{v} = \overline{\mathbf{u}}$ in the statement of the Lax-Milgram lemma, and conclude by the coercivity of the bilinear form, and the Cauchy-Schwarz inequality for the inner product.
9. If the bilinear form $A(\mathbf{u}, \mathbf{v})$ is symmetric, continuous, and coercive, it defines a inner product in $\mathbb{H}$ which is equivalent to the one we already have. Apply the orthogonal projection Theorem 2.11 to the space

$$(\mathbb{H}, \langle\cdot,\cdot\rangle_A),$$

as well as the Riesz-Freéchet representation theorem Proposition 2.14. The minimization property is nothing but the corresponding minimization of the distance with this new inner product (Theorem 2.11 again).
10. Apply Stampachhia's theorem for the choice $\mathbf{K} = \mathbb{H}$ which is a vector space, not just a convex set to see that the inequality in Stampachhia's theorem becomes an equality.
11. If $u'_j \rightharpoonup u'$, and v'_j are the corresponding solutions, check that there is some feasible v with $v'_j \rightharpoonup v'$, and then both limits u and v are related through the same differential problem. Since in this situation $v_j \to v$ in L^∞, the only requirement on the integrand F is its continuity.
12. (a) Through a process limit, one can assume that the probability measure μ is finitely-supported

$$\mu = \sum_i \mu_i \delta_{\mathbf{x}_i}, \quad \mu_i > 0, \sum_i \mu_i = 1, \mathbf{x}_i \in \mathbb{X}.$$

In this case, the vector

$$\left(\int_{\mathbb{X}} v_j(\mathbf{x})\,d\mu(\mathbf{x})\right)_j = \left(\sum_i \mu_i v_j(\mathbf{x}_i)\right)_j = \sum_i \mu_i (v_j(\mathbf{x}_i))_j$$

belongs to D, directly from convexity.

(b) This is easily seen through the approximation procedure of the previous item, and the definition of convexity.
(c) It suffices to use the convexity of the function $u(r) = -\log r$.

13. In general terms, convexity of the functional I can be written in the form

$$\int_{J\times J} f(t_1u_1(x) + t_2u_2(x), t_1u_1(y) + t_2u_2(y))\,dx\,dy \le$$
$$t_1 \int_{J\times J} f(u_1(x), u_1(y))\,dx\,dy + t_2 \int_{J\times J} f(u_2(x), u_2(y))\,dx\,dy$$

for arbitrary $t_1, t_2 \ge 0$, $t_1 + t_2 = 1$, and functions u_1, u_2. It is not easy to find a simple necessary and sufficient condition for this inequality to hold in general. For instance, if functions u_1 and u_2 are taken to be constant, then a necessary condition is that the function of one variable

$$u \mapsto f(u, u)$$

must be convex, but this is far from being sufficient because it does not provide any information on f off the diagonal. A bit of experimentation with simple functions can help us realize that there is no simple answer.

14. (a) It suffices to take $W(x, y, \cdot)$ convex for every pair $(x, y) \in (0, 1)^2$, and Φ, increasing.
(b) If there is an explicit dependence of $W(x, y, u, v)$ on the variable u, then there is no way to derive any information of the behavior on the integrals

$$\int_0^1 W(x, y, u_j(x), u_j(y))\,dy$$

when u_j weakly converges to u. Note how this is not an issue in the previous item.

15. One weak lower semicontinuity statement is the following. Assume each $f_i(x, u)$ is convex in u, and F is non-increasing in the following sense

$$F(\mathbf{x}) \le F(\mathbf{y}) \text{ when } x_i \le y_i \text{ for all } i.$$

Then functional I is weak lower semicontinuity. It can be generalized to include a more general statement as follows. Suppose the function $(-1)^i f_i(x, u)$ is convex for each i, and

$$F(\mathbf{x}) \le F(\mathbf{y}), \quad (-1)^i x_i \le (-1)^i y_i.$$

Then functional I is weak lower semicontinuous.

16. (a) This is elementary since

$$
\begin{aligned}
N s_N(f)^2 &= \sum_{i=1}^{N} \left(\int_{\Omega} f(\boldsymbol{x}) u_i(\boldsymbol{x})\, d\boldsymbol{x} \right)^2 \\
&= \sum_{i=1}^{N} \langle f, u_i \rangle^2 .
\end{aligned}
$$

(b) From the last identity and the classical Cauchy-Schwarz inequality, it is clear that

$$s_N(f) \le \|f\|,$$

and so each s_N is continuous. Moreover, because

$$f = \sum_i \langle f, u_i \rangle u_i ,$$

we can deduce that

$$\|f\|^2 = \sum_i \langle f, u_i \rangle^2 = \lim_{N \to \infty} s_N(f),$$

and this limit is monotone. This suffices to finish.

17. This is the final part of Exercise 30 of Chap. 2. The derivative $E'(u)$ can be given the explicit form

$$
\begin{aligned}
E'(u)(x) = &\int_0^x [\psi'(u'(y)) - (x - y)\phi'(u(y))]\, dy \\
&- x \int_0^1 [\psi'(u'(y)) - (1 - y)\phi'(u(y))]\, dy .
\end{aligned}
$$

A.4 Chapter 4

1. This is the infinite-dimensional analogue of Exercise 4 of Chap. 3. The argument is formally the same as the finite-dimensional case.
2. Write

$$\phi_a''(u) = 12[(u + \frac{a}{4})^2 + \frac{a^2}{16}(a^2 - 1)].$$

If $a^2 - 1 \geq 0$, ϕ_a is convex, and the unique minimizer of the problem is the linear function $u(t) = \alpha t$. If $a^2 - 1 < 0$, there is a part of the graph of ϕ_a that is not convex, and so existence of minimizers is compromised. Take $a = 1/2$, and explore the corresponding variational problem, to check that there is a small interval J in such a way that if $\alpha \notin J$, the lineal function still is the unique minimizer, but if $\alpha \in J$, then there are infinitely many minimizers, but the linear function is not a minimizer any longer.

3. Check that the functional admits minimizers because convexity dependence with respect to u' is strictly convex. Minimizers are solutions of the differential equation $u'' = u^2$ which are convex functions of one variable.
4. In the proof of Theorem 4.2, there is a step where a certain difference of integrals should be shown to converge to zero (indicated as a proposed exercise: this one). Use the fact that for a.e. $a \in J$, the integrand $F(x, \mathbf{u}, \mathbf{U})$ is bounded from above by some constant when triplets $(x, \mathbf{u}, \mathbf{U})$ belong to a box

$$[a - \delta/2, a + \delta/2] \times J_{\mathbf{u}} \times J_{\mathbf{U}}$$

where $J_{\mathbf{u}}$ and $J_{\mathbf{U}}$ are specific compact sets for the variables $\mathbf{u}$ and $\mathbf{U}$, respectively. Use the dominated convergence theorem to conclude that such difference of integrals tend to zero.
5. (a) This is an elementary Multivariate Calculus exercise.
   (b) This has already been shown in the final section of the chapter.
6. Missing technical details refer to the justification of differentiation under the integral sign. This is correct under the bound assumed on the integrand F and its partial derivatives.
7. It is not difficult to conclude that if $m(\alpha)$ is the value of the infimum, then $m(\alpha) = 1$ for $\alpha = 0$ and $\alpha \geq 1$, while $m(\alpha) = 0$, else.
8. Consider perturbations in the subspace

$$H^1_{per}(J; \mathbb{R}^n) = \{\mathbf{v} \in H^1(J; \mathbb{R}^n) : \mathbf{v}(x_0) - \mathbf{v}(x_1) = \mathbf{0}\},$$

to conclude that (4.20) is valid for all $\mathbf{v}$ in this subspace. Proceed in two successive steps:

   (a) take first, in particular, $\mathbf{v} \in H^1_0(J; \mathbb{R}^N)$ to conclude (4.21) as in Theorem (4.4);
   (b) for $\mathbf{v} \in H^1_{per}(J; \mathbb{R}^n)$ such that the constant vector

$$\mathbf{v}_0 = \mathbf{v}(x_0) = \mathbf{v}(x_1)$$

   is free, and bearing in mind (4.21), go back to (4.20) and, by integrating by parts, conclude that this time the end-point contributions amount to

$$F_{\mathbf{U}}(x, \mathbf{u}(x), \mathbf{u}'(x))\Big|_{x_0}^{x_1} \cdot \mathbf{v}_0 = 0.$$

The arbitrariness of $\mathbf{v}_0$ leads to

$$F_{\mathbf{U}}(x, \mathbf{u}(x), \mathbf{u}'(x))\Big|_{x_0}^{x_1} = \mathbf{0}.$$

This final condition, together with differential system (4.21) (and the periodicity conditions for $\mathbf{u}$), make up the optimality problem.

9. Mimicking the process around the proof of Theorem 4.21, one finds that

$$F_{\mathbf{u}}(x, \mathbf{u}(x), \mathbf{u}'(x), \mathbf{u}''(x)) - \frac{d}{dx} F_{\mathbf{U}}(x, \mathbf{u}(x), \mathbf{u}'(x), \mathbf{u}''(x))$$
$$+ \frac{d^2}{d^2x} F_{\mathbf{Z}}(x, \mathbf{u}(x), \mathbf{u}'(x), \mathbf{u}''(x)) = \mathbf{0} \text{ a.e. } x \text{ in } J,$$

if $F(x, \mathbf{u}, \mathbf{U}, \mathbf{Z})$ is the integrand for the second-order functional

$$\int_J F(x, \mathbf{u}(x), \mathbf{u}'(x), \mathbf{u}''(x))\, dx$$

under end-point conditions both for $\mathbf{u}$ and $\mathbf{u}'$.

10. Perform an integration by parts on the second term in (4.20) to find that

$$\int_J [F_{\mathbf{u}}(x, \mathbf{u}(x), \mathbf{u}'(x)) - \frac{d}{dx} F_{\mathbf{U}}(x, \mathbf{u}(x), \mathbf{u}'(x))] \cdot \mathbf{v}(x)\, dx = 0$$

because contributions from end-points vanish. The arbitrariness of $\mathbf{v}$, other than end-points restrictions, imply (4.21).

11. (a) Corollary 4.1 can be applied to deduce the existence of solutions for each fixed $\epsilon > 0$.
    (b) The Euler-Lagrange system becomes

$$-\epsilon u_1'' + (u_1 u_2 - 1)u_2 = 0, \quad -\epsilon u_2'' + (u_1 u_2 - 1)u_1 = 0,$$

    which is a singularly-perturbed, second-order, non-linear differential system.
    (c) The previous system is impossible to solve explicitly. However, its asymptotic limit as ϵ becomes smaller and smaller can be figured out. In fact, in the complement of a smaller and smaller symmetric interval with respect to $t = 1/2$, the solution will essentially follow the curve $u_1 u_2 - 1 = 0$. In that particular, central subinterval due to symmetry, the curve will follow momentarily the principal diagonal of the plane.
12. Use the same ideas as in Exercise 8 above.
13. Neumann boundary conditions cannot be imposed directly in a first-order system because competing functions will, typically have weak derivatives in a certain Lebesgue space, and these cannot admit individual point values. Instead,

consider the problem for the modified functional

$$\int_J \left(\frac{1}{2}u'(x)^2 + F(x, u(x))\right) dx - u_0' u(0) + u_1' u(1)$$

minimized among all feasible u's in the appropriate space without any end-point conditions.

14. This is the typical situation with a saw-tooth sequence of functions where only slopes ± 1 are used, and they alternate in smaller and smaller scales, while preserving the given end-point conditions. All this functions belong to $\mathcal{A}$, and yet its weak limit, which is the trivial function, does not belong to $\mathcal{A}$ because the integral condition would yield 1 instead of $\sqrt{2}$.
15. Checking that (4.27) is a solution of (4.26) is an interesting Calculus exercise. On the other hand, it is elementary to argue that problem (4.26) can admit at most one solution.
16. This exercise involves an interesting manipulations to find a way to perform a first integration in the Euler-Lagrange system. The statement itself yields the clue.
17. (a) In order to apply our main existence theorem for such a variational problem, the main point to discuss is the coercivity. Note that the strict convexity of $\boldsymbol{f}$ in the derivative is very easily checked. If we put $\boldsymbol{f} = \boldsymbol{f}(\mathbf{y})$,

$$0 \le \left(\frac{1}{2}|\mathbf{z}| - |\boldsymbol{f}|\right)^2 \le \frac{1}{4}|\mathbf{z}|^2 - |\mathbf{z}|\,|\boldsymbol{f}| + |\boldsymbol{f}|^2,$$

and

$$|\mathbf{z}|\,|\boldsymbol{f}| \le \frac{1}{4}|\mathbf{z}|^2 + |\boldsymbol{f}|^2.$$

In this way, since

$$\frac{1}{2}|\mathbf{z} - \boldsymbol{f}|^2 \ge \frac{1}{2}|\mathbf{z}|^2 - |\mathbf{z}|\,|\boldsymbol{f}|,$$

we can conclude that

$$\frac{1}{2}|\mathbf{z} - \boldsymbol{f}|^2 \ge \frac{1}{4}|\mathbf{z}|^2 - |\boldsymbol{f}(\mathbf{y})|^2. \tag{A.1}$$

This is not exactly the coercivity condition required because our lower bound incorporates an explicit dependence on the variable $\mathbf{y}$. However, for arbitrary $t \in J = [0, T]$, we have

$$|\mathbf{x}(t) - \mathbf{x}_0|^2 \le t \int_0^t |\mathbf{x}'(s)|^2\, ds,$$

due to Jensen's inequality

$$\left|\frac{1}{t}\int_0^t \mathbf{x}'(s)\,ds\right|^2 \le \frac{1}{t}\int_0^t |\mathbf{x}'(s)|^2\,ds.$$

Then, by (A.1),

$$\begin{aligned}|\mathbf{x}(t)-\mathbf{x}_0|^2 \le& t\int_0^t |\mathbf{x}'(s)|^2\,ds\\ &\le 4T\,E(\mathbf{x}) + 4T\int_0^t |\boldsymbol{f}(\mathbf{x}(s))|^2\,ds.\end{aligned}$$

For the last term, in turn, we find that

$$\begin{aligned}\int_0^t |\boldsymbol{f}(\mathbf{x}(s))|^2\,ds \le& 2\int_0^t |\boldsymbol{f}(\mathbf{x}(s))-\boldsymbol{f}(\mathbf{x}_0)|^2\,ds + 2\int_0^t |\boldsymbol{f}(\mathbf{x}_0)|^2\,ds\\ &\le 2M^2\int_0^t |\mathbf{x}(s)-\mathbf{x}_0|^2\,ds + 2T\,\boldsymbol{f}(\mathbf{x}_0)|^2.\end{aligned}$$

Altogether, we obtain

$$\begin{aligned}|\mathbf{x}(t)-\mathbf{x}_0|^2 \le\ & 4T\,E(\mathbf{x})\\ &+ 8M^2T\int_0^t |\mathbf{x}(s)-\mathbf{x}_0|^2\,ds\\ &+ 8T^2|\boldsymbol{f}(\mathbf{x}_0)|^2.\end{aligned}$$

If $E(\mathbf{x})$ belongs to a bounded set of numbers, we can conclude, through Gromwall's lemma, that $|\mathbf{x}(t)-\mathbf{x}_0|$ is uniformly bounded, and in this case (A.1) becomes the appropriate necessary coercivity. More specifically, suppose $\{\mathbf{x}_j\}$ is minimizing for E. In this case, the sequence of numbers $\{E(\mathbf{x}_j)\}$ is certainly bounded, and by our hypotheses and our discussion above, $\{\mathbf{x}_j\}$ is bounded uniformly:

$$\mathbf{x}_j(t) \in \mathbf{K} \text{ for all } j \text{ and all } t \in J,$$

for a certain, fixed compact set $\mathbf{K}$. Once we can count on this piece of information, (A.1) implies

$$\begin{aligned}4\int_J |\mathbf{x}_j'(s)|^2\,ds \le& E(\mathbf{x}_j) + \int_J \max_{\mathbf{K}} |\boldsymbol{f}(\mathbf{y})|^2\,ds\\ \le& E(\mathbf{x}_j)\\ &+ \int_J [2M^2 \max_{\mathbf{K}} |\mathbf{y}-\mathbf{x}_0|^2 + 2|\boldsymbol{f}(\mathbf{x}_0)|^2]\,ds < \infty,\end{aligned}$$

and so $\{\mathbf{x}_j\}$ is uniformly bounded in $H^1(J;\mathbb{R}^n)$. This is the main conclusion of coercivity.

(b) The Euler-Lagrange system can be rewritten in the form

$$\mathbf{e}(t)' + \nabla \boldsymbol{f}(\mathbf{x}(t))^T \mathbf{e}(t) = \mathbf{0} \text{ in } J, \quad \mathbf{e}(T) = \mathbf{0},$$

where $\mathbf{x}(t)$ is any minimizer of the problem, and

$$\mathbf{e}(t) = \mathbf{x}'(t) - \boldsymbol{f}(\mathbf{x}(t))$$

is the residual associated with such minimizer. Conclude that the only solution for the previous linear problem for $\mathbf{e}$ is the trivial one.

18. The situation for a constant vector $\mathbf{y}$ or a variable path $\mathbf{y}(t)$ is the same. The existence of an optimal path is shown, under the conditions given for the map $\boldsymbol{f}(\mathbf{x},\mathbf{y})$, by checking how the uniformity of the Lipschitz constant M permits to exactly reproduce the proof in Exercise 17, both in the constant and variable cases. Note that the problem is first-order in $\mathbf{x}$, but zero-order in $\mathbf{y}$. Optimality conditions become a differential-algebraic system of the form

$$\mathbf{e}'(t) + \nabla_{\mathbf{x}} \boldsymbol{f}(\mathbf{x}(t),\mathbf{y}(t))^T \mathbf{e}(t) = \mathbf{0} \text{ in } (0,T),$$
$$\nabla_{\mathbf{y}} \boldsymbol{f}(\mathbf{x}(t),\mathbf{y}(t))^T \mathbf{e}(t) = \mathbf{0} \text{ in } (0,T),$$

if

$$\mathbf{e}(t) = \mathbf{x}'(t) - \boldsymbol{f}(\mathbf{x}(t),\mathbf{y}(t))$$

is the residual vector. One would like to be able to conclude, from these optimality conditions, that $\mathbf{e} \equiv \mathbf{0}$. However, it is not easy to give an explicit answer unless we assume a more explicit form of $\boldsymbol{f}$. In the case of a linear mapping

$$\boldsymbol{f}(\mathbf{x}, y) = \mathbf{A}\mathbf{x} + \mathbf{b}y,$$

those optimality conditions become

$$\mathbf{e}' + \mathbf{A}^T \mathbf{e} = \mathbf{0}, \quad \mathbf{b}^T \mathbf{e} = 0.$$

From here one can deduce that $\mathbf{e} \equiv 0$ if and only if the rank condition of Kalman

$$\operatorname{rank}\left(\mathbf{b}\ \mathbf{A}\mathbf{b}\ \mathbf{A}^2\mathbf{b} \ldots \mathbf{A}^{n-1}\mathbf{b}\right) = n$$

is verified.

19. In this case, one considers apparently more general variations as indicated in the statement. Consider the real function

$$s \in (-\epsilon, \epsilon) \mapsto g(s) \equiv \int_0^1 F(t, \mathbf{u}(s,t), \mathbf{u}_t(s,t))\, dt,$$

and impose the condition that $g'(0) = 0$. Though one is considering more general variations, optimality is reduced to the usual Euler-Lagrange system of optimality.

20. Regarding the path $\mathbf{u}$ as fixed, we perform a change of variables setting

$$\mathbf{v}(t) = \mathbf{u}(\phi^{-1}(t)),$$

and consider the variational problem for ϕ

$$\text{Minimize in } \phi : \quad \int_0^1 F(\mathbf{u}_\phi(t), \mathbf{u}'_\phi(t))\, dt$$

under the constraints

$$\phi(0) = 0, \phi(1) = 1, \quad \phi' > 0.$$

The functional to be minimized can be explicitly written, through the change of variables $s = \phi^{-1}(t)$, as

$$\int_0^1 F\left(\mathbf{u}(s), \frac{1}{\phi'(s)}\mathbf{u}'(s)\right) \phi'(s)\, ds.$$

It is with respect to this integral as a functional of ϕ, for fixed $\mathbf{u}$, that one should write optimality conditions.

21. This is a continuation of the previous exercise to stress how inner-variations are especially well-suited to study optimality under point-wise constraints. Note how paths of the form

$$\mathbf{u}_\psi(t) = \mathbf{u}(\psi(t))$$

comply with point-wise constraint if **u** itself does for arbitrary function ψ. In particular, if **u** is optimal, then one can explore optimality by looking at the real function

$$s \in (-\epsilon, \epsilon) \mapsto g(s) = \mathbf{u}(t + s\psi(t)), \quad \psi(0) = 0, \psi(1) = 0,$$

and asking for $g'(0) = 0$.

22. This new situation does not deserve special comments. It is a matter of following the formalism for optimality conditions by considering the sections

$$s \in (-\epsilon, \epsilon) \mapsto g(s) = F\left(\int_J f(x, u(x) + sv(x), u'(x) + sv'(x))\, dx\right)$$

for fixed v with $v(1) = v(0) = 0$, if u is indeed a minimizer.

23. For this problem we have the integrand

$$F(\mathbf{x}, \mathbf{u}) = [\mathbf{u}^T \mathbf{A}(\mathbf{x})\mathbf{u}]^{1/2}.$$

It is not easy to write the corresponding Euler-Lagrange system in a meaningful, compact form. In fact, to identify the Chrystoffel symbols, it is better to derive the system in a fully developed form

$$F(\mathbf{x}, \mathbf{u}) = (\sum_{i,j} A_{ij}(\mathbf{x}) u_i u_j)^{1/2}, \quad \mathbf{A} = (A_{ij}), \mathbf{u} = (u_i).$$

For the two-dimensional case, we have

$$F(\mathbf{x}, \mathbf{u}) = \sqrt{A_{11}(x_1, x_2)u_1^2 + 2A_{12}(x_1, x_2)u_1 u_2 + A_{22}(x_1, x_2)u_2^2}.$$

It is easier to write the two-equation system for this particular case.

24. Through optimality conditions for the minimization problem in ψ for fixed **u**, it is easy to find that

$$I_i(\mathbf{u}) = \left(\int_0^1 |\mathbf{u}'(t)|\, dt\right)^2,$$

which is not a local integral functional. The calculation for the second situation are trivial and one finds that $I = I_i$.

25. The application of basic results in this chapter to the problem proposed for fixed ϵ is straightforward (though it requires some care in calculations). The basic property to define the limit problem is to write the weak limit of the quotiens $1/a_\epsilon(x)$ in the same quotient form

$$\frac{1}{a_\epsilon(x)} \rightharpoonup \frac{1}{a(x)}.$$

A.5 Chapter 5

1. This is a straightforward, practice exercise.
2. Use linearity to translate the given fact about balls into the claimed statment: the map $\mathbf{T}$ is open.
3. This is just a direct argument.
4. Under the condition on $\mathbf{T}$, it is standard to check that the series determining $\mathbf{S}$ is well-defined. Then check easily that

$$(\mathbf{1}-\mathbf{T})\mathbf{S}=\mathbf{S}(\mathbf{1}-\mathbf{T})$$

   is the identity.
5. (a) This is straightforward.
   (b) This amounts to checking that $\mathbf{T}$ is injective but not onto.
   (c) $\mathbf{T}'f(x)=f(x/2)$.
6. (a) To show that π is linear, argue as follows. For $\mathbf{x},\mathbf{y}\in\mathbb{H}$ arbitrary, and a scalar α, we have

$$\langle\pi(\alpha\mathbf{x}),\mathbf{y}\rangle=\langle\alpha\mathbf{x},\pi\mathbf{y}\rangle=\langle\alpha\mathbf{x},\pi^2\mathbf{y}\rangle$$

   and then

$$\langle\pi(\alpha\mathbf{x}),\mathbf{y}\rangle=\alpha\langle\mathbf{x},\pi^2\mathbf{y}\rangle=\cdots=\alpha\langle\pi\mathbf{x},\mathbf{y}\rangle.$$

   The arbitrariness of $\mathbf{y}$ implies $\pi(\alpha\mathbf{x})=\alpha\pi\mathbf{x}$, and in a similar way for sums of vectors. For the continuity, bear in mind that

$$\|\pi\mathbf{x}\|^2=\langle\pi^2\mathbf{x},\mathbf{x}\rangle=\langle\pi\mathbf{x},\mathbf{x}\rangle\leq\|\pi\mathbf{x}\|\,\|\mathbf{x}\|.$$

   This inequality also shows that $\|\pi\|\leq 1$. But notice that

$$\|\pi^2\mathbf{x}\|=\|\pi\mathbf{x}\|,$$

   and so $\|\pi\|=1$.
   (b) Identify the four statements as (a), (b), (c), and (d), respectively. Then (a) means that

$$\langle\mathbf{x}-\pi\mathbf{x},\pi\mathbf{x}\rangle=0,\quad\langle\mathbf{x},\pi\mathbf{x}\rangle=\|\pi\mathbf{x}\|^2\geq 0,$$

   which is (b). (c) trivially implies (d).
7. Since $\|\mathbf{Tx}\|=\|\mathbf{x}\|$, $\|\mathbf{T}\|=1$, and so

$$\sigma(\mathbf{T})\subset\{|\lambda|\leq 1\}.$$

To show the equality, for arbitrary λ with $|\lambda| \leq 1$ prove that $(\mathbf{T} - \lambda\mathbf{1})$ is not onto by trying to find the inverse image of, for instance, $(1, 0, \dots)$.

8. It is clear that $e(\mathbf{T}) = \{1/j : j \in \mathbb{N}\}$. Since $\|\mathbf{T}\| = 1$,
$$\sigma(\mathbf{T}) \subset \{|\lambda| \leq 1\}. \tag{A.2}$$
If λ has size not greater than unity, it is non-zero, and not one of the $1/n$, then
$$(\mathbf{T} - \lambda\mathbf{1})\mathbf{x} = \mathbf{y}, \quad \mathbf{x}_j = \frac{1}{1/j - \lambda}\mathbf{y}_j.$$
For j large, $\|\lambda - 1/j| \geq |\lambda|/2$, and then vector $\mathbf{x}$ given by the above formula belongs to ℓ^p. Operator $\mathbf{T} - \lambda\mathbf{1}$ is bijective, and by the open-mapping theorem is isomorphism. This implies that (A.2) is an equality.

9. (a) The linearity and continuity is straightforward. For the composition, use Fubini's theorem to show that
$$\mathbf{T}^2 u(x) = \int_0^x K_2(x, y) u(y)\, dy$$
where
$$K^2(x, y) = \int_y^x K(x, z) K(z, y)\, dz.$$
Keep iterating. In fact, if
$$\mathbf{T}u(x) = \int_0^x K(x, y) u(y)\, dy, \quad \mathbf{S}u(x) = \int_0^x L(x, y) u(y)\, dy,$$
then
$$\mathbf{T}(\mathbf{S}u)(x) = \int_0^x \overline{K}(x, y) u(y)\, dy$$
where
$$\overline{K}(x, y) = \int_y^x K(x, z) L(z, y)\, dz.$$
(b) There is no special difficulty with induction based on the previous calculations.

(c) A constant kernel have a trivial kernel for the corresponding operator. The kernel

$$K(x, y) = x \cos\left(\frac{\pi}{x} y\right)$$

has a constant function in the kernel of the corresponding operator.

(d) In the case $K(x, y) = 1$, the rule found in the first item leads to

$$K_j(x, y) = \frac{1}{j!}(x - y)^j, \quad j \geq 0.$$

In this way, one can write

$$\begin{aligned}(\mathbf{1} - \mathbf{T})^{-1}u(x) &= \sum_j \mathbf{T}^j u(x) \\ &= \sum_j \int_0^x \frac{1}{j!}(x - y)^j u(y)\, dy \\ &= \int_0^x e^{x-y} u(y)\, dy. \end{aligned} \tag{A.3}$$

(e) This is straightforward.

10. (a) Apply Corollary 5.1 to conclude that the limit operator

$$\mathbf{S}\mathbf{y} = \sum_j \mathbf{T}^j \mathbf{y}$$

is linear and continuous, and because $(\mathbf{1} - \mathbf{T})\mathbf{S} = \mathbf{1}$, $\mathbf{S}$ is the inverse of $(\mathbf{1} - \mathbf{T})$ which is, therefore an isomorphism of $\mathbb{E}$.

(b) Show that it is a Cauchy sequence based on the convergence of the sum of powers of $\mathbf{T}$.

(c) Write

$$\mathbf{x}_j = \sum_{k=0}^{j-1} \mathbf{T}^k \mathbf{y} + \mathbf{T}^j \mathbf{x}_0, \quad \mathbf{x} = \sum_{k=0}^{\infty} \mathbf{T}^k \mathbf{y}.$$

The difference $\mathbf{x} - \mathbf{x}_j$ becomes

$$\sum_{k=j}^{\infty} \mathbf{T}^k \mathbf{y} - \mathbf{T}^j \mathbf{x}_0 = \mathbf{T}^j \left((\mathbf{1} - \mathbf{T})^{-1} \mathbf{y} - \mathbf{x}_0\right).$$

11. Suppose $\{u'_j\}$ is an infinite number of linearly independent C^0-functions in $[0, 1]$, which we can assume uniformly bounded but with the same size in terms of the sup-norm. This effect can be achieved by multiplying each derivative by an appropriate constant. Then we know that, at least for a subsequence and subtracting a fixed function, $u'_j \rightharpoonup 0$ but this convergence cannot be strong (because the norm is the same and away from zero). Then $u_j \to 0$ (strong), but their derivatives converge only weak, and so the derivative operator cannot be continuous.
12. It is clear that $\mathbf{T}$ is linear and continuous. If a number λ belongs to the resolvent then $\mathbf{T} - \lambda\mathbf{1}$ must be a bijection:

    (a) $\mathbf{T} - \lambda\mathbf{1}$ is injective: all level sets $\{a = \lambda\}$ of a must be negligible;
    (b) $\mathbf{T} - \lambda\mathbf{1}$ is onto: the function $1/(a(\boldsymbol{x}) - \lambda)$ must belong to $L^\infty(\Omega)$.

    Eigenvalues are those λ's for which the level set $\{a = \lambda\}$ is non-negligible, while λ belongs to the resolvent when

$$\frac{1}{a(\boldsymbol{x}) - \lambda} \in L^\infty(\Omega).$$

13. Use the change of variables $st = r$ to see that

$$\lim_{s\to\infty} \mathbb{L}(u)(s) = 0.$$

    Through the mean-value theorem for the exponential, and the same change of variables prove that $\mathbb{L}u$ is a continuous function, and that $\mathbb{L}$ is continuous. The property of the derivative of $\mathbb{L}u$ is standard.
14. Define a linear operator $\mathbf{T}$ and its adjoint $\mathbf{T}'$, through the Riesz representation theorem, by putting

$$A(\mathbf{u}, \mathbf{v}) = \langle \mathbf{T}\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{T}'\mathbf{v}\rangle.$$

    Check, through the coercivity of A and the Cauchy-Schwarz inequality, that

$$c\|\mathbf{v}\| \le \|\mathbf{T}'\mathbf{v}\|.$$

    The duality relations imply that the closure of $\mathbf{R}(\mathbf{T})$ is the full space $\mathbb{H}$. Conclude by checking that $\mathbf{T}$ is closed.
15. This is the same operator of the two final items of Exercise 9. It does not have eigenvalues, and its norm is $e - 1$.

A.6 Chapter 6

1. This is immediate from the definitions upon some reflection.
2. Let $\mathbf{u} \in \mathbb{U}$. By hypothesis,

$$\mathbf{u} = \mathbf{m} + (1/2)\mathbf{u}_1, \quad \mathbf{m} \in \mathbb{M}, \mathbf{u}_1 \in \mathbf{U}.$$

By the same reason, there are

$$\mathbf{m}_1 \in \mathbb{M}, \mathbf{u}_2 \in \mathbf{U}, \quad \mathbf{u}_1 = \mathbf{m}_1 + \frac{1}{2}\mathbf{u}_2$$

i.e.

$$\mathbf{u} = \mathbf{m} + \frac{1}{2}(\mathbf{m}_1 + \frac{1}{2}\mathbf{u}_2) = \mathbf{m} + \frac{1}{2}\mathbf{m}_1 + \frac{1}{2^2}\mathbf{u}_2.$$

Proceed in a similar manner to find that

$$\mathbf{u} = \sum_{i=0}^{n} \frac{1}{2^i}\mathbf{m}_i + \frac{1}{2^{n+1}}\mathbf{u}_{n+1}, \quad \mathbf{m} = \mathbf{m}_0,$$

for every n. This decomposition of $\mathbf{u}$ implies, because $\mathbb{U}$ is bounded, that $\mathbf{u} \in \overline{\mathbb{M}}$.

3. The orthogonality relations are immediate upon interpretation of them in a Hilbert space scenario. For the equality of spectra, it suffices to check that the resolvents are the same. Suppose that $(\mathbf{T}' - \lambda\mathbf{1})\mathbf{x} = \mathbf{0}$, then one concludes that

$$\langle \mathbf{x}, (\mathbf{T} - \lambda\mathbf{1})\mathbf{y} \rangle = 0$$

for all $\mathbf{y}$. If $\mathbf{T} - \lambda\mathbf{1}$ is a bijection, then this can only happen if $\mathbf{x} = \mathbf{0}$, and $(\mathbf{T}' - \lambda\mathbf{1})$ is injective. Note that

$$(\mathbf{T} - \lambda\mathbf{1})' = \mathbf{T}' - \lambda\mathbf{1},$$

and use the orthogonality relations to conclude.

4. (a) If $\mathbf{A}$ is a symmetric, positive-matrix, it has non-negative eigenvalues $\lambda_j > 0$, and it is diagonalizable

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}, \quad \mathbf{D} = \mathrm{diag}(\lambda_j), \mathbf{P}, \text{ non-singular.}$$

We can define, then, in a coherent way

$$\log \mathbf{A} = \mathbf{P} \log \mathbf{D}\mathbf{P}^{-1}, \quad \log \mathbf{D} = \mathrm{diag}(\log \lambda_j).$$

For the concrete matrix provided, it is elementary to find that

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & -2 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} -1 & -1 & -1 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix},$$

and

$$\log \mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & -2 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \log 2 \end{pmatrix} \begin{pmatrix} -1 & -1 & -1 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix},$$

i.e.

$$\log \mathbf{A} = \log 2 \begin{pmatrix} 2 & 1 & 1 \\ -4 & -2 & -2 \\ 2 & 1 & 1 \end{pmatrix}.$$

It is easy to check that $\exp(\log \mathbf{A}) = \mathbf{A}$.

(b) The formula

$$\mathbf{T}u(t) = t \int_0^1 u(s)(1-s)\,ds - \int_0^t u(s)(t-s)\,ds$$

yields an explicit form to the operator $\mathbf{T}$ which proves that $\mathbf{T}$ is compact. However, it is better, for other purposes, to use its definition in the statement. Two successive integration by parts show that

$$\int_U'' V\,ds = \int_U V''\,ds, \quad U(0) = U(1) = V(0) = V(1),$$

and hence $\mathbf{T}$ is self-adjoint. In addition, one integration by parts leads to

$$\langle \mathbf{T}u, u\rangle = \int_0^1 u'(s)^2\,ds \geq 0,$$

and $\mathbf{T}$ is positive as well. Consequently, for any continuous function $f(r) : [0, \infty) \to [0, \infty)$, the operator $f(\mathbf{T})$ can be defined as in Example 6.8.

5. This exercise reduces to checking that the operator

$$\mathbf{T} : L^2(0, 1) \mapsto L^2(0, 1), \quad U = \mathbf{T}u$$

defined through

$$-U''(x) + U(x) = u(x) \text{ in } (0, 1), \quad U'(0) = U'(1) = 0, \tag{A.4}$$

is compact, self-adjoint, and positive, just as in the second part of the previous exercise. There is no explicit formula for $\mathbf{T}$ though.

(a) In the first place $\mathbf{T}$ is well-defined. Note that problem A.4 admits a unique solution for $u \in L^2(0, 1)$ as it is the minimizer of the quadratic functional

$$\int_0^1 [\frac{1}{2}U'(x)^2 + \frac{1}{2}U(x)^2 - u(x)U(x)]\,dx$$

under no boundary condition. Refer to Corollary 4.2.

(b) For the compactness, multiply (A.4) by U itself and integrate by parts to find

$$\int_0^1 [U'(x)^2 + U(x)^2]\,dx = \int_0^1 u(x)U(x)\,dx.$$

From here, we deduce that

$$\|U\| \le \|u\| \text{ in } L^2(0, 1),$$
$$\|U'\|^2 \le \|u\|\,\|U\| \le \|u\|^2 \text{ in } L^2(0, 1).$$

Hence, if u belongs to a bounded set in $L^2(0, 1)$, so does the set of derivatives U', and by Proposition 2.7, the set of functions U themselves belong to a compact set in $L^2(0, 1)$.

6. Since by definition, it is always true that

$$\mathbf{1} = \pi_i + \pi_i^\perp, \quad \pi_i \to \mathbf{1}$$

we should have the given statement. However, if it were true that $\|\pi_i^\perp\| \to 0$, that would imply that the identity operator $\mathbf{1}$ would be compact, which is impossible in an infinite-dimensional Hilbert space. There is no contradiction with the Banach-Steinhaus principle.

7. Because $\pi_i \to \mathbf{1}$ point-wise, if $\mathbf{T}$ is compact then $\mathbf{T}(\mathbf{B})$ is a compact set and hence

$$\|\pi_i \mathbf{T} - \mathbf{T}\| = \sup_{\mathbf{x}:\|\mathbf{x}\|=1} |(\pi_i \mathbf{T} - \mathbf{T})\mathbf{x}| \to 0, \quad i \to \infty.$$

Conversely if the norm in the statement tends to zero, $\mathbf{T}$ turns out to be a limit, in the operator norm, of a sequence of finite-rank operator, and hence, compact.

8. The arguments are much in the spirit of Proposition 2.13. Bear in mind that under the given hypotheses $\pi_i \pi_j = \mathbf{0}$ for $i \neq j$. For the compactness use the fact that a limit of finite-rank operators is compact.

9. This is a typical Linear Algebra exercise stated in an infinite-dimensional scenario.
10. (a) From

$$\lambda u(x) = \int_0^x (1-x)yu(y)\,dy + \int_x^1 (1-y)xu(y)\,dy,$$

it is easy to check the first item by differentiating twice with respect to x.

(b) This is elementary from the previous item. The eigen-functions are

$$u_j(x) = \sin(\pi j x).$$

(c) For integral operators of this kind, the adjoint is the integral operator corresponding to the symmetric kernel

$$K'(x,y) = K(y,x) = \begin{cases} (1-y)x, & 0 \le x \le y \le 1, \\ (1-x)y, & 0 \le y \le x \le 1, \end{cases} = K(x,y).$$

$\mathbf{T}$ is then self-adjoint. The Fredholm alternative ensures that the equation

$$(\mathbf{T} - (j\pi)^{-2}\mathbf{1})u = v$$

is solvable, provided $\langle v, w\rangle = 0$ for all w such that

$$(\mathbf{T} - (j\pi)^{-2}\mathbf{1})w = 0,$$

i.e. v belongs to the orthogonal complement of the eigen-space associated with $(\pi j)^{-2}$.

11. The one in Exercise 7 is not compact, for example, by Theorem 6.3: every compact operator must have eigenvalues. However, the one in Exercise 8 is compact. To argue this, consider the finite-rank operator

$$\mathbf{T}_j\mathbf{x} = (x_1, x_2/2, x_3/3, \dots, x_j/j, 0, 0, \dots, 0, \dots),$$

and realize that

$$\|\mathbf{T} - \mathbf{T}_j\| \le \frac{1}{j+1}.$$

12. If f is non-constant, its image contains a certain non-negligible interval $J \subset [0,1]$. The sequence $u_j(x) = \sin(jf(x))$ is continuous, uniformly bounded, but it will not converge in the subinterval $f^{-1}(J)$.

13. Since $\mathbf{T}$ is compact and self-adjoint, the Fredholm alternative ensures the result if the unique solution $u(x)$ of the homogeneous equation

$$u(x) - \int_0^\pi \sin(x+y)u(y)\,dy = 0$$

is the trivial one. Write this integral equation in the form

$$\begin{aligned} u(x) =& \sin x \int_0^\pi \cos y u(y)\,dy + \cos x \int_0^\pi \sin y u(y)\,dy \\ =& a \sin x + b \cos x, \end{aligned}$$

and write $u(y) = a \sin y + b \cos y$ back into

$$a = \int_0^\pi \sin y(a \sin y + b \cos y)\,dy,$$

$$b = \int_0^\pi \cos y(a \sin y + b \cos y)\,dy,$$

to conclude that $a = b = 0$.

14. It is very natural to argue that such a $\mathbf{T}$ is a limit of a sequence of finite-rank operators. For a counterexample, check examples of the form given in Exercise 8 of Chap. 5 that has been discussed in Exercise 11 above.
15. In a reflexive Banach space, the unit ball $\mathbf{B}$ is weakly compact. This suffices to prove the claim easily. In the explicit situation given, the sequence $\mathbf{T}u_j$ converges, in the sup norm, to the function 0 for $x \leq 0$; x for $x \geq 0$, whose derivative is discontinuous, and hence, it cannot belong to the image of $\mathbf{T}$.
16. This property has sufficiently emphasized earlier.
17. This proof is a very typical application of Riesz lemma. If the space is not finite-dimensional, a sequence can be found inductively in the unit sphere that cannot converge to anything. The argument is similar to the proof of Proposition 6.3.
18. (a) To calculate $\|\mathbf{T}\|$, evaluate $\mathbf{T}(f_r)$ for

$$f_r = \chi_{[r,1]}, \quad 0 \leq r \leq 1.$$

(b) For any $\lambda \in [0, 1]$, it is easy to realize that the equation

$$(s - \lambda) f(s) = g(s)$$

for f, g in $L^2(0, 1)$ is impossible. However, there is no particular difficulty if $\lambda \notin [0, 1]$. The operator does not have eigenvalues, and so it cannot be compact (Theorem 6.3).

19. (a) This is just a simple calculation.

(b) Both terms in $\mathbf{T}f$ represent compact, integral operators of Volterra type in both cases.

(c) The operator derivative

$$f(x) \mapsto f(x) - \int_0^1 f(y)\, dy$$

is, however, not compact because the first term is just the identity, but the second contribution is compact.

A.7 Chapter 7

1. This is elementary. Take the open set $\pi_N \Omega$, and select a test function ψ for such an open set in $\mathbb{R}^{N-1}$ with $\psi(\boldsymbol{x}_0') = 1$.
2. This is also a geometric property which is, at least, easy to visualize.
3. This is just a change-of-variables issue.
4. There is nothing surprising in this exercise.
5. (a) If $\psi \in C_c^\infty(\mathbb{R}^N)$ and $\phi \in C_c^\infty(\Omega)$, the product $\psi\phi$ is a smooth function with compact support in Ω, and hence

$$\int_\Omega u \nabla(\psi\phi)\, d\boldsymbol{x} = -\int_\Omega \nabla u\, \psi\phi\, d\boldsymbol{x}.$$

Thanks to the product rule for smooth functions,

$$\int_\Omega u \nabla(\psi\phi)\, d\boldsymbol{x} = \int_\Omega u(\nabla\psi\phi + \nabla\phi\psi)\, d\boldsymbol{x}.$$

If we reorganize these identities we find

$$\int_\Omega u\psi\nabla\phi\, d\boldsymbol{x} = -\int_\Omega (\nabla u\psi + u\nabla\psi)\phi\, d\boldsymbol{x}.$$

The arbitrariness of ϕ leads to the desired result.

(b) This is straightforward from the definition.

6. This can be accomplished through the same ideas as in Example 2.12 of Chap. 2. Apply the one-dimensional construction in that example to the function

$$\chi(\boldsymbol{x}) = \chi_t(\boldsymbol{x} \cdot \boldsymbol{n})$$

where χ_t is the 1-periodic function of one variable of that example, and $\boldsymbol{n}$ is any unit normal vector.

7. (a) Just as in the case of the standard Sobolev spaces, one can define the Hilbert space

$$L^2_{div}(\Omega) = \{\mathbf{F} \in L^2(\Omega; \mathbb{R}^N) : \operatorname{div} \mathbf{F} \in L^2(\Omega)\}$$

where the weak divergence $\operatorname{div} \mathbf{F}$ is defined through the usual integration-by-parts formula

$$\int_\Omega \operatorname{div} \mathbf{F} \phi \, d\boldsymbol{x} = -\int_\Omega \mathbf{F} \cdot \nabla \phi \, d\boldsymbol{x}$$

valid for every test function ϕ in Ω.

(b) This is straightforward because $\operatorname{div} \mathbf{F} = 0$ means, according to the previous formula,

$$\int_\Omega \mathbf{F} \cdot \nabla \phi \, d\boldsymbol{x} = 0$$

for all test functions. These are dense in $H^1_0(\Omega)$.

8. This exercise is left for exploration to interested readers.
9. This is similar to arguments in Sects. 2.10 and 7.3.
10. Use Corollary 7.1 and Lemma 7.5 for an increasing sequence of balls $\{\mathbf{B}_j\}$ with radii tending to ∞ to conclude that the set $\{C^\infty_c(\mathbb{R}^N)\}$ is dense in $W^{1,p}(\mathbb{R}^N)$.
11. These facts are straightforward.
12. This is an interesting exercise but straightforward after the proof of Lemma 2.4.
13. It is easy to calculate that if a function $u(x_1, x_2) \in H^1(\mathbf{B})$ is independent of x_2, then

$$\|\nabla u\|^2_{L^2(\mathbf{B};\mathbb{R}^2)} = 2\int_{-1}^1 u'(x_1)^2 \sqrt{1 - x_1^2} \, dx_1.$$

Similarly

$$\|u\|^2_{L^2(\mathbf{B})} = 2\int_{-1}^1 u(x_1)^2 \sqrt{1 - x_1^2} \, dx_1.$$

In this way, one can define the Sobolev space with weight

$$L^2_w(-1, 1) = \{u(x), \text{ measurable} : \int_{-1}^1 u(x)^2 w(x) \, dx < \infty\},$$

$$H^1_w(-1, 1) = \{u(x) \in L^2_w(-1, 1) : u'(x) \in L^2_w(-1, 1)\},$$

for the weight function

$$w(x) = \sqrt{1 - x^2}, \quad x \in (-1, 1).$$

For a general domain the procedure is formally the same for the weight function

$$w(x) = |\{x_1 = x\}|.$$

This procedure can be implemented in a completely general way.

14. That $\mathbb{L}$ is a subspace is straightforward. It is also easy to argue that it is a subspace of $H^1_w(-1, 1)$ with the ideas and notation of the previous exercise. To check that it is not closed, it suffices to find a function in $H^1_w(-1, 1)$ which does not belong to $\mathbb{L}$. The function

$$u(x) = \int_0^x \frac{1}{\sqrt[4]{1 - t^2}}\, dt$$

is a good candidate.

A.8 Chapter 8

1. By looking with care at the integration by parts involved in deriving the weak form of the Euler-Lagrange equation in 8.7, one can conclude that

$$F_{\mathbf{u}}(u(\boldsymbol{x}), \nabla u(\boldsymbol{x})) \cdot \boldsymbol{n}(\boldsymbol{x}) = 0 \text{ on } \partial\Omega. \tag{A.5}$$

Proceed in two steps:

(a) use first a test function $V \in H^1_0(\Omega)$ for which the integration by parts does not interfere with boundary contributions, to derive the PDE itself;
(b) once one can rely on the information coming from the previous PDE, use a general test function $V \in H^1(\Omega)$ to conclude (A.5).

2. This is a direct application of our main existence result. Since the functional is quadratic, it can also be solved with the help of the Lax-Milgram lemma. The optimality condition reads

$$\operatorname{div}(\nabla v - \mathbf{F}) = f \text{ in } \Omega, \quad u = u_0 \text{ on } \partial\Omega.$$

3. The first part is a consequence of the first exercise of this chapter. It amounts to writing the natural boundary condition

$$\mathbf{Au} \cdot \mathbf{n} = 0 \text{ on } \partial\Omega.$$

It is relevant to note the presence of the quadratic term in u to avoid the situation in which minimizing sequences do not remain bounded in $H^1(\Omega)$ by adding an arbitrary constant to u. If one is interested in adjusting a non-vanishing normal derivative, then the functional should be formally changed to

$$\int_\Omega \frac{1}{2}|\nabla u(\boldsymbol{x})|^2\, d\boldsymbol{x} + \int_{\partial\Omega} h(\boldsymbol{x})u(\boldsymbol{x})\, dS(\boldsymbol{x}).$$

4. The integrand for this problem can be writte n in the form

$$\nabla u^T \left(\frac{1}{2}\mathbf{1} + \mathbf{F}\otimes\mathbf{G}\right)\nabla u.$$

   It is therefore a quadratic variational problem and one needs the condition

$$a > 2\|\mathbf{F}\|_{L^\infty(\Omega)}\|\mathbf{G}\|_{L^\infty(\Omega)}$$

   to guarantee convexity and coercivity. The Euler-Lagrange equation is

$$\operatorname{div}\left(a\nabla u + \frac{1}{2}\mathbf{F}\cdot\nabla u\mathbf{G} + \frac{1}{2}\mathbf{G}\cdot\nabla u\mathbf{F}\right) = 0.$$

5. The functional with integrand

$$\mathbf{u} \mapsto \frac{a}{2}|\mathbf{u}|^2 + \frac{b}{2}|\mathbf{u}|\mathbf{F}\cdot\mathbf{u}$$

   has the given Euler-Lagrange equation. The condition

$$a > \|b\mathbf{F}\|_{L^\infty(\Omega)}$$

   is required to ensure existence of solutions.
6. According to the third item of Proposition 8.2, the function f needs to be convex and increasing to ensure convexity of F. Moreover, one needs the asymptotic behavior

$$\lim_{r\to\infty}\frac{f(r)^p}{r} = \alpha > 0, \qquad p > 1.$$

   In fact, it suffices

$$\lim_{r\to\infty}\frac{f(r)}{r} = +\infty.$$

7. This is similar to Exercise 29 of Chap. 2.

8. If the field **F** is divergence-free, then it this term does not affect the underlying Euler-Lagrange equation. On the one hand, by the divergence theorem

$$\int_{\Omega} \nabla u(\boldsymbol{x}) \cdot \mathbf{F}(\boldsymbol{x})\, d\boldsymbol{x} = \int_{\partial\Omega} u(\boldsymbol{x})\mathbf{F}(\boldsymbol{x}) \cdot \mathbf{n}(\boldsymbol{x})\, dS(\boldsymbol{x}),$$

and this last integral is independent of u as it only depends on its boundary values which are fixed. This is a null-lagrangian, meaning that this second term does not affect the optimization process, as it is constant. On the other, it is elementary to realize that when writing the Euler-Lagrange equation, this term drops out precisely because **F** is divergence-free.
9. This is straightforward.
10. (a) The application of the Lax-Milgram lemma is direct in the space $\mathbb{H}$.
    (b) We have to change $\mathbb{H}$ slightly to incorporate the normal boundary condition. In this case the subspace of $L^2(\Omega; \mathbb{R}^N)$ made up of gradients of functions in $H^1(\Omega)$ is the orthogonal complement to $\mathbb{H}$. Hence, optimality becomes

$$\operatorname{curl}(\mathbf{AF} + \mathbf{G}) = \mathbf{0}.$$

Therefor there exists u with $\nabla u = \mathbf{AF} + \mathbf{G}$. This is formally equivalent, by solving for **F**, to

$$\operatorname{div}\left[\mathbf{A}^{-1}(\nabla u - \mathbf{G})\right] = 0 \text{ in } \Omega,$$
$$\mathbf{A}^{-1}(\nabla u - \mathbf{G}) \cdot \mathbf{n} = f \text{ on } \partial\Omega.$$

According to Exercise 3 of this same chapter, these are exactly the optimality condition for the standard problem for the functional

$$\int_{\Omega} \left(\frac{1}{2}\nabla u^T \mathbf{A}^{-1} \nabla u - \nabla u^T \mathbf{A}^{-1}\mathbf{G}\right) d\boldsymbol{x} + \int_{\partial\Omega} uf\, dS(\boldsymbol{x})$$

over $H^1(\Omega)$.
11. (a) The standard theorem cannot be applied because the matrix

$$\mathbf{A} = \frac{1}{2}\begin{pmatrix} 1 & a \\ a & a^2 \end{pmatrix}$$

is not positive definite.
    (b) The proposed perturbation transforms the previous matrix **A** to

$$\overline{\mathbf{A}} = \frac{1}{2}\begin{pmatrix} 1 & a \\ a & a^2 + \epsilon \end{pmatrix}.$$

This time the theorem can be applied and there is a unique solution.

(c) The situation is exactly the same. The main change for this new problem would affect the form of the Euler-Lagrange equation that it would be non-linear.

12. This is a straightforward generalization of Exercise 1. The corresponding optimality problem would read

$$-\operatorname{div}[F_{\mathbf{u}}(u, \nabla u)] + F_u(u, \nabla u) = 0 \text{ in } \Omega,$$
$$u = u_0 \text{ on } \Gamma,\ F_{\mathbf{u}}(u, \nabla u) \cdot \mathbf{n} = 0 \text{ on } \partial\Omega \setminus \Gamma.$$

13. This is a zeroth-order problem. A typical existence theorem would require the convexity of the integrand $F(\boldsymbol{x}, \cdot)$ for a.e. $\boldsymbol{x} \in \Omega$, and the coercivity

$$C(|u|^p - 1) \le F(\boldsymbol{x}, u), \quad C > 0,\ p > 1.$$

Optimality conditions would amount to

$$F_u(\boldsymbol{x}, u_\lambda(\boldsymbol{x})) = \lambda, \text{ constant in } \Omega,$$

where such constant value λ is determined so that the integral constraint is respected

$$\int_\Omega u_\lambda(\boldsymbol{x})\, d\boldsymbol{x} = |\Omega| u_0.$$

14. It is well-known from Vector Calculus courses that the flux of a vector field $\mathbf{F}$ across a given surface $\mathbb{S}$ is given by the surface integral

$$\int_{\mathbb{S}} \mathbf{F}(\boldsymbol{x}) \cdot \mathbf{n}(\boldsymbol{x})\, dS(\boldsymbol{x}),$$

where $\mathbf{n}(\boldsymbol{x})$ is the unit normal vector field to $\mathbb{S}$ (one of the possible two options). If surface $\mathbb{S}$ is given by the graph of a certain function $u(x, y)$ over Ω, then the flux is

$$\int_\Omega \mathbf{F}(x, y, u(x, y)) \cdot (-u_x(x, y), -u_y(x, y), 1)\, dx\, dy.$$

More explicitly, if we put

$$\mathbf{F}(x, y, u(x, y)) = (F_1(x, y, u(x, y)), F_2(x, y, u(x, y)), F_3(x, y, u(x, y))),$$

the flux is given by

$$\int_\Omega [F_3(x, y, u(x, y)) - F_1(x, y, u(x, y))u_x(x, y) \\ - F_2(x, y, u(x, y))u_y(x, y)]\, dx\, dy.$$

After a few careful calculation, the equation of optimality is written

$$\operatorname{div} \mathbf{F}(x, y, u(x, y)) = 0,$$

which forces optimal surfaces to have their image on the set where the divergence of the vector field $\mathbf{F}$ vanishes. This is clearly impossible if this possibility is forbidden by the function u_0 furnishing the boundary values for u. If $\mathbf{F}$ is divergence-free, the Euler-Lagrange equation becomes void, and the integrand becomes a null-lagrangian.

15. (a) It is easy to check that the square of the determinant is not convex. For instance, the section

$$t \mapsto F\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + t\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right).$$

is not convex.

(b) With a bit of care the equation becomes

$$\nabla(\det \nabla^2 u) \cdot (u_{yy} - u_{xy}, u_{xx} - u_{xy}) = 0 \text{ in } \Omega,$$

which is a highly non-linear equation.

16. (a) It is immediate to show the existence of a unique solution either through the Lax-Milgram theorem (we are talking about a quadratic functional) or by our main existence result in the chapter.

(b) Optimality conditions lead to

$$\int_\Omega [\alpha_1 \chi(\boldsymbol{x}) + \alpha_0(1 - \chi(\boldsymbol{x}))]\nabla u(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x})\, d\boldsymbol{x} = 0$$

for all $v \in H_0^1(\Omega)$. Use successively test functions v with support in Ω_1 and Ω_0 to conclude that u must be harmonic separately in both sets. Once this information is available, for a general v with arbitrary values on Γ, use the divergence theorem and deduce the transmission condition

$$(\alpha_1 - \alpha_0)\nabla u \cdot \mathbf{n} = 0 \text{ on } \Gamma.$$

Both α's could be functions of $\boldsymbol{x}$.

17. (a) Once again, it is immediate to show existence of a unique solution for each ϵ fixed, either through the Lax-Milgram lemma or our main existence result for variational problems.
    (b) Argue that the sequence of minimizers $\{u_\epsilon\}$ is uniformly bounded in $H^1(\Omega)$, and, hence, possibly for a subsequence (not relabeled) it converges weakly to some limit u.
    (c) Under the hypothesis suggested one would find that

$$0 = \int_\Omega a_\epsilon \nabla v_\epsilon \cdot \nabla u_\epsilon \, d\boldsymbol{x} \to \int_\Omega a_0 \nabla v \cdot \nabla u \, d\boldsymbol{x}.$$

Hence

$$\int_\Omega a_0 \nabla v \cdot \nabla u \, d\boldsymbol{x} = 0$$

for all $v \in H_0^1(\Omega)$, and this implies that u_0 is the minimizer of the limit functional.

18. (a) This is again a consequence of our main existence theorem for convex, coercive variational problems.
    (b) As in the mentioned exercise, the transmission condition would be

$$[|\nabla \hat{u} - \nabla u|^{p-2}(\nabla \hat{u} - \nabla u) - |\nabla \hat{u}|^{p-2} \nabla \hat{u}] \cdot \mathbf{n} = 0 \text{ on } \partial\Omega.$$

    (c) Because the value of the functional for U is finite (it is important that $U = u$ in Ω), independently of ϵ, we can deduce that $\hat{u}_\epsilon \to u$ in Ω.

19. (a) This is standard after main results in this chapter.
    (b) This is a particular case of the same question in the previous exercise. The main difference is the linearity of the operation taking each u to the corresponding minimizer v_ϵ.
    (c) Note that

$$m_\epsilon = \frac{1}{2} \|\nabla v_\epsilon\|^2_{L^2(\mathbb{R}^N \setminus \Omega)} + \frac{1}{2\epsilon} \|\nabla v_\epsilon - \nabla u\|^2_{L^2(\Omega)}, \tag{A.6}$$

and bear in mind that we can take

$$\|v_\epsilon\|^2_{H^1(\mathbb{R}^N \setminus \Omega)} = \|\nabla v_\epsilon\|^2_{L^2(\mathbb{R}^N \setminus \Omega)}, \quad \|v_\epsilon\|^2_{H^1(\mathbb{R}^N)} = \|\nabla v_\epsilon\|^2_{L^2(\mathbb{R}^N)}.$$

Then

$$\|v_\epsilon\|^2_{H^1(\mathbb{R}^N)} = \|v_\epsilon\|^2_{H^1(\mathbb{R}^N \setminus \Omega)} + \|\nabla v_\epsilon\|^2_{L^2(\Omega)},$$

and, on the one hand, by (A.6),

$$\|v_\epsilon\|^2_{H^1(\mathbb{R}^N\setminus\Omega)} \le 2m_\epsilon,$$

whereas on the other,

$$\begin{aligned}\|\nabla v_\epsilon\|^2_{L^2(\Omega)} &\le \left(\|\nabla v_\epsilon - \nabla u\|_{L^2(\Omega)} + \|\nabla u\|_{L^2(\Omega)}\right)^2\\ &\le 2\|\nabla v_\epsilon - \nabla u\|^2_{L^2(\Omega)} + 2\|\nabla u\|^2_{L^2(\Omega)}\\ &\le 4m_\epsilon\epsilon + 2\|\nabla u\|^2_{L^2(\Omega)}.\end{aligned}$$

Since we are interested in small values of ϵ, these computations show that there is a uniform constant C such that

$$\|\mathcal{E}_\epsilon u\| \le C\|u\|.$$

(d) Conclude accordingly.

20. The natural possibility would be to consider the subspace

$$\mathbb{H} = \{u \in C^\infty(\Omega) : \mathbf{F}\cdot\nabla u = 0 \text{ on } \partial\Omega\},$$

of $H^1(\Omega)$, and its closure in $H^1(\Omega)$, which we designate by the same letter $\mathbb{H}$. However, it is not difficult to realize that $\mathbb{H}$ becomes the full $H^1(\Omega)$ because one can modify a given function $u \in H^1(\Omega)$ in a small amount (in the $H^1(\Omega)$-norm) near de boundary to make it comply with the given boundary condition. This is similar to the situation of the $L^2(\Omega)$-closure of the set $C_c^\infty(\Omega)$ of smooth functions with compact support in Ω, which is the full $L^2(\Omega)$. The problem is not well-posed. It is ill-posed.

21. This is a practice exercise with typical integrands in the quadratic case. It amounts to checking (strict) convexity and coercivity in each particular case, as well as smoothness conditions to write the underlying Euler-Lagrange equation, at least formally. We briefly comment on each case.

(a) The integrand is

$$F(u_1, u_2) = \frac{1}{2}(|u_1| + |u_2|)^2.$$

It is strictly convex and coercive.

(b) The integrand is quadratic corresponding to the symmetric, positive definite matrix

$$\mathbf{A} = \begin{pmatrix} 2 & -1\\ -1 & 1\end{pmatrix}.$$

(c) Similar to the previous one for the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Though it is convex (not strict), it is not coercive. Check the diagonal $u_1 = u_2$.

(d) The integrand is

$$F(u_1, u_2) = \frac{1}{2}(u_1^4 + u_2^4)^{1/2}.$$

It is strictly convex and coercive. Note that

$$F(\mathbf{u}) = \frac{1}{2}\|\mathbf{u}\|_4^2, \quad \mathbf{u} = (u_1, u_2), \|\mathbf{u}\|_4^4 = u_1^4 + u_2^4.$$

All norms are equivalent in $\mathbb{R}^2$, and the p-th norm is strictly convex for $p > 1$.

(e) The integrand is

$$F(u_1, u_2) = \frac{1}{2}\left(u_1^2\ u_2^2\right)\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} u_1^2 \\ u_2^2 \end{pmatrix}^{1/2}.$$

It is strictly convex and coercive.

(f) The integrand is

$$F(u_1, u_2) = \frac{1}{2}\left((u_1^4 + u_2^2)^{1/2} + u_2^2\right).$$

It is strictly convex and coercive.

(g) Slight variation of the previous one.

(h) The integrand is

$$F(u_1, u_2) = \frac{1}{2}(u_1^2 + u_2^2) + 7\exp(-(u_1 - 1)^4 - u_2^2).$$

It is coercive but not convex.

(i) This time

$$F(u_1, u_2) = |u_1|\,|u_2|.$$

It is neither convex nor coercive.

(j) The integrand is

$$F(u_1, u_2) = (1 + u_1^2)^{1/2}(1 + u_2^2)^{1/2}.$$

It has linear growth at infinity; it is not enough for coercivity. It is not convex either.

(k) This time

$$F(u_1, u_2) = (1 + u_1^2 + u_2^2)^{1/2}.$$

It has linear growth, but it is strictly convex.

22. (a) The mixed boundary condition can be implemented by minimizing the functional over the set of feasible functions

$$\{u \in H^1(\Omega) : u = u_0 \text{ on } \Gamma_0\}.$$

It requires to examine the subspace of $H^1(\Omega)$, given by

$$\{u \in H^1(\Omega) : u = 0 \text{ on } \Gamma_0\},$$

as the closure in $H^1(\Omega)$ of the same subspace of smooth functions.

(b) Mimic the non-homogeneous Neumman boundary condition by minimizing the augmented functional

$$\frac{1}{2}\int_\Omega |\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x} + \frac{\gamma}{2}\int_{\partial\Omega} u^2(\boldsymbol{x}) \, dS(\boldsymbol{x}).$$

(c) If boundary condition is restricted through an inequality, then the part of Dirichlet condition will be the coincidence set $u = u_0$ on $\partial\Omega$, while the Neumman part $\nabla u \cdot \mathbf{n} = 0$ will correspond to its complement where $u < u_0$ on $\partial\Omega$.

23. This is a practice exercise. With a bit of care, one finds that the corresponding Euler-Lagrange equation turns out to be

$$\operatorname{div}\left[(\mathbf{1} - \nabla w(\boldsymbol{x}) \otimes \nabla w(\boldsymbol{x}))\, \nabla u(\boldsymbol{x})\right] = 0.$$

A.9 Chapter 9

1. Almost all main ingredients of the computations have been given in the statement. The formula in the first part is just a careful use of the chain rule. Note that, unless we differentiate twice with respect to the normal to $\partial\Omega$, i.e. with respect to X_N, at least one derivative has to be computed tangentially, and so it must vanish. In order to conclude (9.24), recall that we learn from Differential Geometry that the quantity

$$\nabla X_N \nabla^2 X_N \nabla X_N$$

yields the curvature H of a hyper-surface if it is given by a level set of X_N, and ∇X_N is unitary.

2. The optimality condition for the suggested constrained variational problem is

$$\Delta u = \lambda u \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega,$$

for a positive multiplier λ. The smallest possible such λ is the first eigenvalue of the Laplacian. From this equation one concludes that

$$\lambda \int_\Omega u(\boldsymbol{x})^2 \, d\boldsymbol{x} = \int_\Omega |\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x}.$$

On the other hand, the best constant C in Poincaré's inequality will be such that

$$\int_\Omega u(\boldsymbol{x})^2 \, d\boldsymbol{x} \le C \int_\Omega |\nabla u(\boldsymbol{x})|^2 \, d\boldsymbol{x}.$$

3. For the case $N = 2$, the exponent $p = (N+2)/(N-2)$ cannot be taken directly, so one needs to work with an arbitrary p, apply the corresponding inequalities, and then take p to ∞.
4. This is an elementary Calculus exercise. For $x \ge R$, write

$$f'(x) = \frac{p}{x} f(x) + g(x), \quad g(x) \ge 0.$$

Use the explicit formula

$$f(x) = \exp\left(\int_R^x \frac{p}{s}\, ds\right)\left[f(R) + \int_R^x \exp\left(-\int_R^y \frac{p}{s}\, ds\right) g(y)\, dy\right]$$

for the solution of such a linear, first-order ODE to conclude, based on the fact $g(x) \ge 0$, that

$$f(x) \ge \frac{f(R)}{R} x^p, \quad x \ge R.$$

From here, it is easy to deduce the final result.

5. The main point is to define the operator $\mathbf{T}$ in this new setting appropriately. Put $U = \mathbf{T}u$ for the unique minimizer of the functional

$$\int_\Omega \left[\frac{1}{2}|\nabla U(\boldsymbol{x})|^2 - U(\boldsymbol{x})u(\boldsymbol{x})\right] d\boldsymbol{x}$$

over the space $H^1(\Omega) \cap L^2_0(\Omega)$ This constraint on the mean value of functions is necessary to have a substitute for Poincaré's inequality through Wirtinger's

inequality, and to have that the L^2-norm of the gradient is again a norm in $H^1(\Omega) \cap L_0^2(\Omega)$. Everything else is like in the Dirichlet case.

6. (a) This variational problem may not have solutions. Note that if $\{u_j\}$ is a minimizing sequence, their gradients are uniformly bounded in L^2, and there is a weak limit u. But this weak limit may not comply with the constraint (being an equality constraint). In dimension 1, a simple example may help us in realizing the difficulty. The function

$$u_h(x) = \begin{cases} \sqrt{h} - \frac{x}{\sqrt{h}}, & 0 \le x \le \frac{1}{h}, \\ 0, & \frac{1}{h} \le x \le 1, \end{cases}$$

is such that

$$\int_0^1 u_h'(x)^2 \, dx = 1, \quad \int_0^1 u_h(x)^2 \, dx = \frac{h^2}{3}.$$

A minimizer u would be such that

$$\int_0^1 u'(x)^2 \, dx = 1, \quad \int_0^1 u(x)^2 \, dx = 0$$

which is, obviously, impossible.

(b) It is easy to realize that the proposed problem is equivalent to

$$\text{Minimize in } u \in H^1(\Omega) : \quad -\int_\Omega |u(\mathbf{x})|^2 \, d\mathbf{x}$$

subject to

$$\int_\Omega |\nabla u(\mathbf{x})|^2 \, d\mathbf{x} \le 1.$$

This time the direct method furnishes minimizers, because we have an integral constraint in the derivative in the form of an inequality.

7. According to Theorem 9.13, the Laplace operator Δ (under Dirichlet boundary conditions) has a sequence of positive eigenvalues λ_j and eigenfunctions u_j, which make up an orthonormal basis of $L^2(\Omega)$. For a $L^2(\Omega)$-function u, write the decomposition

$$u = \sum_j \langle u, u_j \rangle u_j$$

and define

$$f(\Delta)u = \sum \langle u, u_j \rangle f(\lambda_j) u_j.$$

8. No comment. This is a practice exercise.
9. No comment. This is a practice exercise.
10. No comment. This is a practice exercise.
11. It is a matter of computing the derivative

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} \int_\Omega \frac{1}{2} |\nabla u(\boldsymbol{x} + \epsilon \Phi(\boldsymbol{x}))(\mathbf{1} + \epsilon \nabla \Phi(\boldsymbol{x}))|^2 \, d\boldsymbol{x}.$$

With a bit of care, one finds that

$$0 = \int_\Omega \Big[\nabla u(\boldsymbol{x}) \nabla^2 u(\boldsymbol{x}) \Phi(\boldsymbol{x}) + \nabla u(\boldsymbol{x}) \cdot \nabla \Phi(\boldsymbol{x}) \, \nabla u(\boldsymbol{x}) \Big] \, d\boldsymbol{x}$$

for all such Φ. After rewriting this integral and using an integration by parts, we have

$$0 = \int_\Omega [\Big[\nabla u(\boldsymbol{x}) \nabla^2 u(\boldsymbol{x}) \Phi(\boldsymbol{x}) - \operatorname{div}[\nabla u(\boldsymbol{x}) \otimes \nabla u(\boldsymbol{x})] \Phi(\boldsymbol{x}) \Big] \, d\boldsymbol{x}.$$

The arbitrariness of Φ leads to

$$\nabla u \nabla^2 u - \operatorname{div}(\nabla u \otimes \nabla u) = \mathbf{0}.$$

Of course, this system reduces, after some calculations, to $\Delta u = 0$.

12. (a) Define the space $H_{div}(\Omega)$ as the subspace of fields $\mathbf{F}(\boldsymbol{x}) : \Omega \to \mathbb{R}^N$ in $L^2(\Omega; \mathbb{R}^N)$ such that

$$\int_\Omega \mathbf{F}(\boldsymbol{x}) \cdot \nabla \phi(\boldsymbol{x}) \, d\boldsymbol{x} = 0$$

for all smooth functions ϕ in Ω. Note how this definition forces both conditions $\operatorname{div} \mathbf{F} = 0$ in Ω and $\mathbf{F} \cdot \boldsymbol{n} = 0$ on $\partial\Omega$. The norm in this space is just the norm of the ambient space $L^2(\Omega; \mathbb{R}^N)$. It is easy to check that $H_{div}(\Omega)$ is a closed subspace of $L^2(\Omega; \mathbb{R}^N)$, and consequently it is also weakly closed. Recall Theorem 3.5.

(b) If the matrix field $\mathbb{A}(\boldsymbol{x})$ is symmetric and uniformly positive-definite in the sense

$$\mathbf{F}^T \mathbb{A}(\boldsymbol{x}) \mathbf{F} \geq C|\mathbf{F}|^2, \quad C > 0, \mathbf{F} \in \mathbb{R}^N,$$

and the vector field $\mathbf{A} \in L^\infty(\Omega; \mathbb{R}^N)$, then the given functional in the statement is coercive, and strictly convex, and the direct method yields a unique minimizer $\hat{\mathbf{F}} \in H_{div}(\Omega)$ of the problem.

(c) To derive the corresponding optimality conditions, we perform variations of the form

$$\hat{\mathbf{F}} + \epsilon\mathbf{F}, \quad \mathbf{F} \in H_{div}(\Omega).$$

The computation

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} E(\hat{\mathbf{F}} + \epsilon\mathbf{F})$$

gives

$$\int_\Omega (\hat{\mathbf{F}}\mathbb{A}\mathbf{F} + \mathbf{A} \cdot \mathbf{F})\, d\boldsymbol{x} = 0$$

for all such $\mathbf{F} \in H_{div}(\Omega)$. According to the definition of $H_{div}(\Omega)$, as the orthogonal complement of gradients of functions in $H^1(\Omega)$ in $L^2(\Omega; \mathbb{R}^N)$, we deduce that

$$\mathbb{A}\hat{\mathbf{F}} + \mathbf{A} = \nabla\phi, \quad \phi \in H^1(\Omega).$$

This can be written, equivalently, in a differential form

$$\operatorname{curl}(\mathbb{A}\hat{\mathbf{F}} + \mathbf{A}) = \mathbf{0} \text{ in } \Omega.$$

Appendix B
So Much to Learn

There is so much material one can look at after having covered the material in this course, depending on preferences and level of understanding, that we have tried to organize it according to the four main areas involved in this textbook: the three occurring in the title, and PDEs. Seeking that these comments may be helpful to students, we have tried to avoid dispersion, and so we highlight a few sources, at most, for each mentioned subject. In some of these, we also provide some clues at our discretion. There is no aim at completeness or relevance on the selection that follows. On the other hand, we have tried to include general references accessible to students; this explains why some more specialized material is not included.

B.1 Variational Methods and Related Fields

This is the broader section in this Appendix, as the Calculus of Variations has been our main interest. Readers will see that, even so, we have left out far too many interesting topics.

B.1.1 Some Additional Sources for the Calculus of Variations

There is a number of fundamental textbooks dealing with various viewpoints on variational methods that can support parts of this text, or go well beyond. We cite the following, some of which are very classic sources, though the goal and methods are not uniform in them [Bu89, Cl13, Da08, Da15, FoLe07, GiMoSo98, Gi03, JoLi98, Mo08, Mor96, Mo03, Ri18].

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4

B.1.2 Introductory Courses

One can feel some curiosity at how variational methods can be introduced even at a more elementary level; or how it is presented to other scientists and/or engineers. We just mention some [BuGiHil98, Da15, Ki18, Ko14, Kot14, Tr96].

B.1.3 Indirect Methods

Though our treatment of variational problems is essentially based on the direct method, there is a lot of fundamental classic material that can be integrated under the umbrella of the indirect method. Inner-variations and corresponding optimality conditions are also relevant in this regard. Some of those issues are

- Legendre-Hadamard condition;
- Weierstrass necessary condition;
- Du Bois-Reymond Lemma;
- Weierstrass-Erdmann conditions.

The encyclopedic work [GiHil96] is a main reference for this area.

B.1.4 Convex and Non-smooth Analysis

Chapter 3 is a mere, brief introduction to Convex Analysis so readers already know a bit what this important area is about; though, again, we have not explored basic material that needs to be studied at some point like, for instance, the theory of conjugate convex functions . In addition to the classical books [EkTe99] or [Ro97], there are excellent additional references like [AtBuMi14, BoLe06, Bo14, Cl90, HiLe01, Mo18, MoNa14, Ta09].

Convex analysis deals, in a general abstract way, with convex sets and convex functions, duality in convex optimization, complementary conditions and problems. Techniques of non-smooth analysis are intimately connected to convex analysis. We mention some important sources for non-smooth analysis in addition to those alreacy cited: [Cl90, Io17].

B.1.5 Lagrangian and Hamiltonian Formalism

Historically, a lot of effort has been devoted to topics within the indirect method. On the other hand, there is a rich tradition on variational methods in Mechanics. Our main reference for this field is [GiHil96], where most of the knowledge in this

area is gathered. This is also a good source for everything related to variations in the independent variable, and Noether's conservation law theorems, as has already been indicated above.

B.1.6 Variational Inequalities

These deal with optimality conditions in optimization problems where the set of competing objects is a convex set of functions or fields but they do not support an underlying linear structure. Two basic references are: [KiSt00, Ul11].

B.1.7 Non-existence and Young Measures

Ever since the pioneering work of Young, Young measures has been a main tool to tame non-convexity and non-existence. Our basic references are [CRV04, FlGo12, Pe97, Yo69].

B.1.8 Optimal Control

Optimal control is a fundamental part of optimization of paramount importance in Engineering. Traditionally is studied together, or after, the Calculus of Variations for one-dimensional problems. Some basic textbooks are: [Be17, Bu14, Le14, Li12, Me09, Tr96, Yo69].

Optimal control is however much more in the sense that the part of this area where state equations or systems are PDE is considerably more complicated, but the scope goes well beyond.

B.1.9 Γ-Convergence

The area of Γ-convergence deals with sequences of functionals and how its limit behavior can be explored in an orderly manner. There are not yet many textbooks in the subject [AtBuMi14, Br02, DM93].

B.1.10 Other Areas

There are a number of important topics that are part of more general fields but they have a character on their own. There are not yet textbooks as they are being intensely explored these days. Some of those are:

- Existence without convexity
- Regularity in variational problems
- Second-order optimality conditions
- Non-local functionals
- Constrained variational problems and multipliers
- Stochastic Calculus of Variations
- Variational problems in L^∞

The two references [Ul11, Is16] treat some of these.

B.2 Partial Differential Equations

The interconnection between the Calculus of Variations and PDEs is so deep that it is impossible to tell them apart. Variational methods are utilized constantly in problems in Analysis where the main interest is the underlying PDEs themselves; and, viceversa, fundamental motivation and techniques in variational problems are constantly borrowed from the field of PDEs. We refer here to additional sources where variational methods are at the background of viewpoints on problems [ACM18, Br13, Br11, CaVi15, CDR18, Cr18, Ev10, GiTr01, Kr08, MaOc19, SVZZ13, Sa16, SBH19, Ta15].

B.2.1 Non-linear PDEs

Non-linear PDEs are sometimes quite different from their linear counterparts, and typically much more difficult. This is almost always true in every part of Analysis. In particular, quite often non-linear PDEs pose quite challenging problems to researchers. Most of the references in the previous item have chapters dealing one way or another with non-linear problems. Some other interesting resources are [AmAr11, Co07].

B.2.2 Regularity for PDEs: Regularity of Ω Is Necessary

The theory of regularity either for variational problems or PDEs is quite delicate and technical, but it should be known to a certain extent by every Applied Analyst. Our basic reference here is [Gr11].

B.2.3 Numerical Approximation

The numerical approximation of solutions to PDEs and variational problems is another fundamental chapter of the theory with a major relevance in applications. The fundamental terms here are finite element analysis. We just name a few books as a guide [Da11, LaBe13, Ma86, Wh17, ZTZ13].

B.3 Sobolev Spaces

The theory of Sobolev spaces is fundamental for Applied Analysis. In this text we have but covered the most basic facts, but a much deeper knowledge is required for a finer treatment of variational problems and PDEs. Some additional references where readers can keep on learning on this area are [Ad03, Ag15, Le17, Ma11, Ta07, Zi89].

B.3.1 Spaces of Bounded Variation, and More General Spaces of Derivatives

Spaces of bounded variation are becoming a cornerstone in many technological applications. They are a first step beyond Sobolev spaces in that derivatives do not necessarily be measurable functions but can be measures. The area is quite mature at this point. Some relevant sources are [AFP00, Li19, Zi89].

There are other important topics that are being intensely investigated these days like fractional Sobolev spaces [RuSi96] or the interplay between PDEs and harmonic analysis [Ab12].

B.4 Functional Analysis

As we have tried to stress in the initial chapters of this text, the role played on Functional Analysis by problems in the Calculus of Variations has been, historically, crucial. Nowadays there is also a clear and important interaction in the other

direction to the extent that one cannot be dispensed with a solid foundation in Functional Analysis to understand the modern Calculus of Variations or the modern theory of PDEs. We name a few additional references for Functional Analysis, in addition to some other more specialized texts in the subsequent subsections [Bo14, Ce10, Ci13, Fa16, LeCl03, Li16, Ov18, Sa17, Si18].

It is again important to stress that Functional Analysis is a very large and fundamental area of Mathematical Analysis that we cannot cover in a few lines, or with a few references.

Nevertheless, we include a few more important subareas of Functional Analysis with some other resources.

- Distributions. The theory of distributions was a fundamental success for Applied Analysis, hardly overestimated ever since [HaTr08, Mi18].
- Unbounded operators and Quantum Mechanics [Sch12].
- Topological vector spaces. Locally convex topological vector spaces [Vo20].
- Orlicz spaces. These spaces are a generalization of Lebesgue spaces that are built by retaining the fundamental properties of the pth-power function that permit that Lebesgue spaces become Banach spaces [HaHa19].
- Non-linear Analysis [AuEk84, Pa18, Ta09].

References

[Ab12] Abels, Hel.: Pseudodifferential and Singular Integral Operators. An Introduction with Applications. De Gruyter Graduate Lectures. De Gruyter, Berlin (2012)

[Ad03] Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. Pure and Applied Mathematics (Amsterdam), vol. 140, 2nd edn. Elsevier, Amsterdam (2003)

[Ag15] Agranovich, M.S.: Sobolev Spaces, Their Generalizations and Elliptic Problems in Smooth and Lipschitz Domains. Revised translation of the 2013 Russian original. Springer Monographs in Mathematics. Springer, Cham (2015)

[AmAr11] Ambrosetti, A., Arcoya, D.: An introduction to nonlinear functional analysis and elliptic problems. In: Progress in Nonlinear Differential Equations and Their Applications, vol. 82. Birkhäuser, Boston (2011)

[AFP00] Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free Discontinuity Problems. Oxford Mathematical Monographs. The Clarendon Press, New York (2000)

[ACM18] Ambrosio, L., Carlotto, A., Massaccesi, A.: Lectures on Elliptic Partial Differential Equations. Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)], vol. 18. Edizioni della Normale, Pisa (2018)

[AtBuMi14] Attouch, H., Buttazzo, G., Michaille, G.: Variational Analysis in Sobolev and BV Spaces. Applications to PDEs and Optimization. MOS-SIAM Series on Optimization, vol. 17, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2014)

[AuEk84] Aubin, J.-P., Ekeland, I.: Applied Nonlinear Analysis. In: Pure and Applied Mathematics (New York). Wiley, New York (1984)

[Be17] Bertsekas, D.P.: Dynamic Programming and Optimal Control, vol. I. 4th edn. Athena Scientific, Belmont (2017)

[BiKr84] Birkhoff, G., Kreyszig, E.: The establishment of functional analysis. Historia Math. **11**, 258–321 (1984)

[BoLe06] Borwein, J.M., Lewis, A.S.: Convex Analysis and Nonlinear Optimization. Theory and Examples. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, vol. 3, 2nd edn. Springer, New York (2006)

[Bo14] Botelho, F.: Functional Analysis and Applied Optimization in Banach Spaces. Applications to Non-convex Variational Models. With contributions by Anderson Ferreira and Alexandre Molter. Springer, Cham (2014)

[Br02] Braides, A.: Γ-Convergence for Beginners. Oxford Lecture Series in Mathematics and its Applications, vol. 22. Oxford University Press, Oxford (2002)

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4

[Br13] Bressan, A.: Lecture Notes on Functional Analysis. With Applications to Linear Partial Differential Equations. Graduate Studies in Mathematics, vol. 143. American Mathematical Society, Providence (2013)

[Br11] Brezis, H.: Functional Analysis, Sobolev Spaces and Partial Differential Equations. Universitext. Springer, New York (2011)

[Bu14] Burns, J.A.: Introduction to the Calculus of Variations and Control—with Modern Applications. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. CRC Press, Boca Raton (2014)

[Bu89] Buttazzo, G.: Semicontinuity, Relaxation and Integral Representation in the Calculus of Variations. Pitman Research Notes in Mathematics Series, vol. 207. Longman Scientific & Technical, Harlow (1989); copublished in the United States with John Wiley, New York

[BuGiHil98] Buttazzo, G., Giaquinta, M., Hildebrandt, S.: One-Dimensional Variational Problems. An Introduction. Oxford Lecture Series in Mathematics and its Applications, vol. 15. The Clarendon Press/Oxford University Press, New York (1998)

[CaVi15] Cañada, A., Villegas, S.: A Variational Approach to Lyapunov Type Inequalities. From ODEs to PDEs. With a foreword by Jean Mawhin. SpringerBriefs in Mathematics. Springer, Cham (2015)

[CRV04] Castaing, C., Raynaud de Fitte, P., Valadier, M.: Young Measures on Topological Spaces. With Applications in Control Theory and Probability Theory. Mathematics and its Applications, vol. 571. Kluwer Academic, Dordrecht (2004)

[Ce10] Cerdà, J.: Linear Functional Analysis. Graduate Studies in Mathematics, vol. 116. American Mathematical Society, Providence, RI; Real Sociedad Matemática Española, Madrid (2010)

[Ci13] Ciarlet, P.G.: Linear and Nonlinear Functional Analysis with Applications. Society for Industrial and Applied Mathematics, Philadelphia (2013)

[CDR18] Cioranescu, D., Donato, P., Roque, M.P.: An Introduction to Second Order Partial Differential Equations. Classical and Variational Solutions. World Scientific, Hackensack (2018)

[Cl90] Clarke, F.H.: Optimization and Nonsmooth Analysis. Classics in Applied Mathematics, vol. 5, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1990)

[Cl13] Clarke, F.: Functional Analysis, Calculus of Variations and Optimal Control. Graduate Texts in Mathematics, vol. 264. Springer, London (2013)

[Co07] Costa, D.G.: An Invitation to Variational Methods in Differential Equations. Birkhäuser, Boston (2007)

[Cr18] Craig, W.: A Course on Partial Differential Equations. Graduate Studies in Mathematics, vol. 197. American Mathematical Society, Providence (2018)

[Da08] Dacorogna, B.: Direct Methods in the Calculus of Variations. Applied Mathematical Sciences, vol. 78, 2nd edn. Springer, New York (2008)

[Da15] Dacorogna, B.: Introduction to the Calculus of Variations, 3rd edn. Imperial College Press, London (2015)

[DM93] Dal Maso, G.: An Introduction to Γ-Convergence. Progress in Nonlinear Differential Equations and Their Applications, vol. 8. Birkhäuser, Boston (1993)

[Da11] Davies, A.J.: The Finite Element Method. An Introduction with Partial Differential Equations, 2nd edn. Oxford University Press, Oxford (2011)

[EkTe99] Ekeland, I., Témam, R.: Convex Analysis and Variational Problems. Translated from the French. Corrected Reprint of the 1976 English Edition. Classics in Applied Mathematics, vol. 28. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1999)

[Ev10] Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, vol. 19, 2nd edn. American Mathematical Society, Providence (2010)

[Fa16] Farenick, D.: Fundamentals of Functional Analysis. Universitext. Springer, Cham (2016)

[FlGo12] Florescu, L.C., Godet-Thobie, C.: Young Measures and Compactness in Measure Spaces. De Gruyter, Berlin (2012)

[FoLe07] Fonseca, I., Leoni, G.: Modern Methods in the Calculus of Variations: Lp Spaces. Springer Monographs in Mathematics. Springer, New York (2007)

[FrGi16] Freguglia, P., Giaquinta, M.: The Early Period of the Calculus of Variations. Birkhäuser/Springer, Cham (2016)

[GiHil96] Giaquinta, M., Hildebrandt, S.: Calculus of Variations. I. The Lagrangian Formalism. II. The Hamiltonian Formalism. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vols. 310–311. Springer, Berlin (1996)

[GiMoSo98] Giaquinta, M., Modica, G., Souček, J.: Cartesian Currents in the Calculus of Variations. I. Cartesian Currents II. Variational Integrals. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vols. 37–38. Springer, Berlin (1998)

[GiTr01] Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Reprint of the 1998 edition. Classics in Mathematics. Springer, Berlin (2001)

[Gi03] Giusti, E.: Direct Methods in the Calculus of Variations. World Scientific, River Edge (2003)

[Go80] Goldstine, H.H.: A History of the Calculus of Variations from the 17th Through the 19th Century. Studies in the History of Mathematics and Physical Sciences, vol. 5. Springer, New York (1980)

[Gr11] Grisvard, P.: Elliptic Problems in Nonsmooth Domains. Reprint of the 1985 original [MR0775683]. With a foreword by Susanne C. Brenner. Classics in Applied Mathematics, vol. 69. Society for Industrial and Applied Mathematics, Philadelphia (2011)

[HaHa19] Harjulehto, P., Hästö, P., Orlicz Spaces and Generalized Orlicz Spaces. Lecture Notes in Mathematics, vol. 2236. Springer, Cham (2019)

[HaTr08] Haroske, D.D., Triebel, H.: Distributions, Sobolev Spaces, Elliptic Equations. EMS Textbooks in Mathematics. European Mathematical Society (EMS), Zürich (2008)

[HiLe01] Hiriart-Urruty, J.B., Lemaréchal, C.: Fundamentals of Convex Analysis. Abridged Version of Convex Analysis and Minimization Algorithms. I [Springer, Berlin, 1993; MR1261420] and II [ibid.; MR1295240]. Grundlehren Text Editions. Springer, Berlin (2001)

[Io17] Ioffe, A.D.: Variational Analysis of Regular Mappings. Theory and Applications. Springer Monographs in Mathematics. Springer, Cham (2017)

[Is16] Ishikawa, Y.: Stochastic Calculus of Variations. For Jump Processes. De Gruyter Studies in Mathematics, vol. 54, 2nd edn. De Gruyter, Berlin (2016)

[JoLi98] Jost, J., Li-Jost, X.: Calculus of Variations. Cambridge Studies in Advanced Mathematics, vol. 64. Cambridge University Press, Cambridge (1998)

[Ki18] Kielhöfer, H.: Calculus of Variations. An Introduction to the One-Dimensional Theory with Examples and Exercises. Translated from the 2010 German original. Texts in Applied Mathematics, vol. 67. Springer, Cham (2018)

[KiSt00] Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. Reprint of the 1980 Original. Classics in Applied Mathematics, vol. 31. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2000)

[Ko14] Komzsik, L.: Applied Calculus of Variations for Engineers, 2nd edn. CRC Press, Boca Raton (2014)

[Kot14] Kot, M.: A First Course in the Calculus of Variations. Student Mathematical Library, vol. 72. American Mathematical Society, Providence (2014)

[Kr94I] Kreyszig, E.: On the calculus of variations and its major influences on the mathematics of the first half of our century. I. Am. Math. Month. **101**(7), 674–678 (1994)

[Kr94II] Kreyszig, E.: On the calculus of variations and its major influences on the mathematics of the first half of our century. II. Amer. Math. Month. **101**(9), 902–908 (1994)

[Kr08] Krylov, N.V.: Lectures on Elliptic and Parabolic Equations in Sobolev Spaces. Graduate Studies in Mathematics, vol. 96. American Mathematical Society, Providence (2008)

[LaBe13] Larson, M.G., Bengzon, F.: The Finite Element Method: Theory, Implementation, and Applications. Texts in Computational Science and Engineering, vol. 10. Springer, Heidelberg (2013)

[LeCl03] Lebedev, L.P., Cloud, M.J.: The Calculus of Variations and Functional Analysis. With Optimal Control and Applications in Mechanics. Series on Stability, Vibration and Control of Systems. Series A: Textbooks, Monographs and Treatises, vol. 12. World Scientific, River Edge (2003)

[Le17] Leoni, G.: A First Course in Sobolev Spaces. Graduate Studies in Mathematics, vol. 181, 2nd edn. American Mathematical Society, Providence (2017)

[Le14] Levi, M.: Classical Mechanics with Calculus of Variations and Optimal Control. An Intuitive Introduction. Student Mathematical Library, vol. 69. American Mathematical Society, Providence; Mathematics Advanced Study Semesters, University Park (2014)

[Li12] Liberzon, D.: Calculus of Variations and Optimal Control Theory. A Concise Introduction. Princeton University Press, Princeton (2012)

[Li16] Limaye, B.V.: Linear Functional Analysis for Scientists and Engineers. Springer, Singapore (2016)

[Li19] Liflyand, E.: Functions of Bounded Variation and Their Fourier Transforms. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, Cham (2019)

[MaOc19] Marin, M., Öchsner, A.: Essentials of Partial Differential Equations. With Applications. Springer, Cham (2019)

[Ma86] Marti, J.T.: Introduction to Sobolev Spaces and Finite Element Solution of Elliptic Boundary Value Problems. Computational Mathematics and Applications. Academic Press [Harcourt Brace Jovanovich, Publishers], London (1986)

[Ma11] Maz'ya, V.: Sobolev Spaces with Applications to Elliptic Partial Differential Equations. Second, revised and augmented edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 342. Springer, Heidelberg (2011)

[Me09] Mesterton-Gibbons, M.: A Primer on the Calculus of Variations and Optimal Control Theory. Student Mathematical Library, vol. 50. American Mathematical Society, Providence (2009)

[Mi18] Mitrea, D.: Distributions, Partial Differential Equations, and Harmonic Analysis. Second edition [MR3114783]. Universitext. Springer, Cham (2018)

[Mo18] Mordukhovich, B.S.: Variational Analysis and Applications. Springer Monographs in Mathematics. Springer, Cham (2018)

[MoNa14] Mordukhovich, B.S., Nam, N.M.: An Easy Path to Convex Analysis and Applications. Synthesis Lectures on Mathematics and Statistics, vol. 14. Morgan & Claypool, Williston (2014)

[Mo08] Morrey, C.B., Jr.: Multiple Integrals in the Calculus of Variations. Reprint of the 1966 edition [MR0202511]. Classics in Mathematics. Springer, Berlin (2008)

[Mor96] Morse, M.: The calculus of Variations in the Large. Reprint of the 1932 original. American Mathematical Society Colloquium Publications, vol. 18. American Mathematical Society, Providence (1996)

[Mo03] Moser, J.: Selected Chapters in the Calculus of Variations. Lecture Notes by Oliver Knill. Lectures in Mathematics ETH Zürich. Birkhäuser, Basel (2003)

[Ov18] Ovchinnikov, S.: Functional Analysis. An Introductory Course. Universitext. Springer, Cham (2018)

[Pa18] Papageorgiou, N.S., Winkert, P.: Applied Nonlinear Functional Analysis. An Introduction. De Gruyter Graduate. De Gruyter, Berlin (2018)

[Pe97] Pedregal, P.: Parametrized Measures and Variational Principles. Progress in Nonlinear Differential Equations and Their Applications, vol. 30. Birkhäuser, Basel (1997)

[Ri18] Rindler, F.: Calculus of Variations. Universitext. Springer, Cham (2018)

[Ro97] Rockafellar, R.T.: Convex Analysis. Reprint of the 1970 Original. Princeton Landmarks in Mathematics. Princeton Paperbacks. Princeton University Press, Princeton (1997)

[RuSi96] Runst, T., Sickel, W.: Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations. De Gruyter Series in Nonlinear Analysis and Applications, vol. 3. Walter de Gruyter, Berlin (1996)

[Sa16] Salsa, S.: Partial Differential Equations in Action. From Modelling to Theory. Unitext, vol. 99, 3rd edn. La Matematica per il 3+2. Springer, Cham (2016)

[SVZZ13] Salsa, S., Vegni, F.M.G., Zaretti, A., Zunino, P.: A Primer on PDEs. Models, Methods, Simulations. Translated and Extended from the 2009 Italian edition. Unitext, vol. 65. La Matematica per il 3+2. Springer, Milan (2013)

[Sa17] Sasane, A.: A Friendly Approach to Functional Analysis. Essential Textbooks in Mathematics. World Scientific, Hackensack (2017)

[SBH19] Sayas, F.J., Brown, T.S., Hassell, M.E.: Variational Techniques for Elliptic Partial Differential Equations. Theoretical Tools and Advanced Applications. CRC Press, Boca Raton (2019)

[Sch12] Schmüdgen, K.: Unbounded Self-Adjoint Operators on Hilbert Space. Graduate Texts in Mathematics, vol. 265. Springer, Dordrecht (2012)

[Si18] Siddiqi, A.H.: Functional Analysis and Applications. Industrial and Applied Mathematics. Springer, Singapore (2018)

[Ta15] Taheri, A.: Function Spaces and Partial Differential Equations. Vol. 1. Classical Analysis. Vol. 2. Contemporary Analysis. Oxford Lecture Series in Mathematics and its Applications, vol. 40–41. Oxford University Press, Oxford (2015)

[Ta09] Takahashi, W.: Introduction to Nonlinear and Convex Analysis. Yokohama Publishers, Yokohama (2009)

[Ta07] Tartar, L.: An Introduction to Sobolev Spaces and Interpolation Spaces. Lecture Notes of the Unione Matematica Italiana, vol. 3. Springer, Berlin; UMI, Bologna (2007)

[Tr96] Troutman, J.L.: Variational Calculus and Optimal Control. With the Assistance of William Hrusa. Optimization with Elementary Convexity. Undergraduate Texts in Mathematics, 2nd edn. Springer, New York (1996)

[Ul11] Ulbrich, M.: Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces. MOS-SIAM Series on Optimization, vol. 11. Society for Industrial and Applied Mathematics (SIAM), Philadelphia; Mathematical Optimization Society, Philadelphia (2011)

[Vo20] Voigt, J.: A Course on Topological Vector Spaces. Compact Textbooks in Mathematics. Birkhäuser, Cham (2020)

[Wh17] Whiteley, J.: Finite Element Methods. A Practical Guide. Mathematical Engineering. Springer, Cham (2017)

[Yo69] Young, L.C.: Lectures on the Calculus of Variations and Optimal Control Theory. Foreword by Wendell H. Fleming W. B. Saunders, Philadelphia (1969)

[Zi89] Ziemer, W.P.: Weakly Differentiable Functions. Sobolev Spaces and Functions of Bounded Variation. Graduate Texts in Mathematics, vol. 120. Springer, New York (1989)

[ZTZ13] Zienkiewicz, O.C., Taylor, R.L., Zhu, J.Z.: The Finite Element Method: Its Basis and Fundamentals, 7th edn. Elsevier/Butterworth Heinemann, Amsterdam (2013)

Index

P. Pedregal, *Functional Analysis, Sobolev Spaces, and Calculus of Variations*,
La Matematica per il 3+2 157, https://doi.org/10.1007/978-3-031-49246-4